

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 . Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: March 17, 2003, 15:54:39 ; Search time 65 Seconds  
(without alignments)  
5104.524 Million cell updates/sec

Title: US-09-763-153-1  
Perfect score: 2337  
Sequence: 1 gccactctctctctggagga.....acaaggagcccccagaag 1245

oring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 60 summaries

Command line parameters:

-MODEL-frame-n2p.model -DEV-rlp  
-DB-A\_Geneseq\_101002 -OFMT-fastan -SUFFIX-n2p.rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi  
-LIST=60 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=20  
-MODE-LOCAL -OUTFMT=ptco -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER-US09763153 @GCL1.1.81 @runat\_11032003.084246.1971 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FPGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=7

Database : A\_Geneseq\_101002.\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	2298	98.3	415	21	Truncated human pr
2	2298	98.3	419	22	Human mature wild
3	2298	98.3	419	22	Wild-type human pr
4	2298	98.3	419	22	Human protein C de
5	2298	98.3	419	23	Human protein C de
6	2298	98.3	460	9	Sequence of human
7	2298	98.3	460	18	Human protein C.
8	2298	98.3	461	7	Sequence of poly
9	2298	98.3	461	8	Human protein C.
10	2298	98.3	461	10	Zymogen form of hu
11	2298	98.3	461	12	Human protein C.
12	2298	98.3	461	12	Human protein C.
13	2298	98.3	461	12	Human protein C.
14	2298	98.3	461	14	Human protein C.
15	2298	98.3	461	17	Human protein C.
16	2298	98.3	461	20	Human protein C.
17	2298	98.3	461	22	Human protein C.
18	2298	98.3	461	22	Human protein C.
19	2298	98.3	461	22	Human protein C.
20	2298	98.3	461	22	Human protein C.
21	2295	98.2	419	23	Human protein C.
22	2295	98.2	419	23	Human protein C.
23	2295	98.2	419	23	Human protein C.
24	2295	98.2	461	9	Human protein C.
25	2294	98.2	461	10	Human protein C.
26	2293	98.1	419	23	Human protein C.
27	2293	98.1	419	23	Human protein C.
28	2293	98.1	419	23	Human protein C.
29	2292	98.1	419	23	Human protein C.
30	2292	98.1	419	23	Human protein C.
31	2292	98.1	419	23	Human protein C.
32	2292	98.1	419	23	Human protein C.
33	2292	98.1	461	12	Human protein C.
34	2292	98.1	461	12	Human protein C.
35	2292	98.1	461	12	Human protein C.
36	2292	98.1	461	12	Human protein C.
37	2291	98.0	419	14	Human protein C.
38	2291	98.0	419	14	Human protein C.
39	2291	98.0	419	23	Human protein C.
40	2291	98.0	419	23	Human protein C.
41	2291	98.0	419	23	Human protein C.
42	2291	98.0	419	23	Human protein C.
43	2290	98.0	419	23	Human protein C.
44	2290	98.0	419	23	Human protein C.
45	2290	98.0	419	23	Human protein C.
46	2290	98.0	419	23	Human protein C.
47	2290	98.0	419	23	Human protein C.
48	2290	98.0	419	23	Human protein C.
49	2290	98.0	419	23	Human protein C.
50	2289	97.9	419	23	Human protein C.
51	2289	97.9	419	23	Human protein C.
52	2289	97.9	419	23	Human protein C.
53	2289	97.9	419	23	Human protein C.
54	2289	97.9	419	23	Human protein C.
55	2289	97.9	419	23	Human protein C.
56	2289	97.9	419	23	Human protein C.
57	2289	97.9	419	23	Human protein C.
58	2289	97.9	419	23	Human protein C.
59	2289	97.9	419	23	Human protein C.
60	2289	97.9	419	23	Human protein C.

## ALIGNMENTS

RESULT 1  
AAV56803  
ID AAV56803 standard; Protein; 415 AA.  
XX



XX Human; protein C derivative; anticoagulation activity; thrombosis;  
 KW serpin inactivation; acute coronary syndrome; myocardial infarction;  
 KW vascular occlusive disorder; hypercoagulable state; angina; sepsis;  
 KW disseminated intravascular coagulation; DIC; burn; transplantation;  
 KW sickle cell disease; viral haemorrhagic fever; protein C deficiency;  
 KW haemolytic uraemic syndrome; acute arterial thrombotic occlusion;  
 KW thromboembolism; prothrombotic disorder; gene therapy; thalassaemia.  
 XX Homo sapiens.  
 XX WO200159084-A1.  
 XX PD 16-AUG-2001.  
 XX 02-FEB-2001; 2001WO-US01221.  
 XX 11-FEB-2000; 2000US-0181948.  
 XX 14-MAR-2000; 2000US-0189199.  
 XX (ELIL) LILLY & CO ELI.  
 XX Gerlitz BE, Grinnell BW, Jones BE;  
 XX WPI; 2001-514662/56.  
 XX DR N-PSDB; RAD15223.  
 XX PT Protein C derivative for treating acute coronary syndromes, vascular  
 PT occlusive disorders, thrombotic disorders and sepsis, comprises  
 PT substitutions at specified amino acid positions -  
 XX Claim 1; Page 43-44; 59pp; English.

CC The invention relates to human protein C derivatives and nucleic acid  
 CC molecules encoding such derivatives. These derivatives have increased  
 CC anticoagulation activity, resistance to serpin inactivation and  
 CC increased sensitivity to thrombin activation compared to wild type  
 CC protein C, and retains the biological activity of the wild type human  
 CC protein C. Protein C derivatives are useful in the manufacture of a  
 CC medicament for the treatment of acute coronary syndromes e.g. myocardial  
 CC infarction and unstable angina; and disease states predisposing to  
 CC thrombosis; vascular occlusive disorders and hypercoagulable states e.g.  
 CC disseminated intravascular coagulation (DIC), burns, transplantations,  
 CC thalassaemia, sickle cell disease, viral haemorrhagic fever and  
 CC haemolytic uraemic syndrome; sepsis in combination with bacterial  
 CC permeability increasing protein; thrombotic disorders in combination  
 CC with an anti-platelet agent; protein C deficiency; acute arterial  
 CC thrombotic occlusion, thromboembolism or stenosis in coronary, cerebral  
 CC or peripheral arteries or in vascular grafts in combination with a  
 CC thrombolytic agent. Nucleic acid molecules of the invention are useful  
 CC for treating humans with genetically predisposed prothrombotic disorders  
 CC by gene therapy. The present sequence is human mature wild type  
 CC protein C.

SQ Sequence 419 AA;

Alignment Scores:  
 Pred. No.: 1,34e-159 Length: 419  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 Gaps: 0

US-09-763-153-1 (1-1245) x AAE08625 (1-419)

QY 1 GCCAACTCTTCGGAGGAGCTCCGTACAGCAGCGCTGGAGCGGAGTGCATAGAGGAG 60  
 Db 1 AlaAsnSerPheLeuGluLeuLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 20  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGGAATTTCCAAATGTGGATGACACACTGGCCTTC 120  
 Db 21 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspPheLeuAlaPhe 40

QY 121 TGTCTCAACGACGCTGACGCTGACCGGTGACCGAGTCTTGGTCTTGGACGACCCCGTGC 180  
 Db 41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60  
 QY 181 AGCTGTGTCTGGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCG 240  
 Db 61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
 QY 241 CGCAGCGCTGGAGGCGCGCTTCTCCAGCGCGGAGGTGAGCTTCCCTCAATGCTCGCTG 300  
 Db 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
 QY 301 GACAAAGCGGCTGCACGCTTACTGCTGAGGAGGTGGGCTGGCGGCTGTAGCTGT 360  
 Db 101 AspAsnGlyGlyCysThrHisTyrcysLeuGluValGlyTrpArgArgCysSerCys 120  
 QY 361 GCGCTGTCTACAAGCTGGGCGGACGCTCTCGAGTGTCCAGTGTCCAGTGTCCAGTGTCC 420  
 Db 121 AlaProGlyTyrcysLeuGlyAspLeuGlnCysHisProAlaValLysPhePro 140  
 QY 421 TGTGGGAGCGCTGGAAGCGGATGAGAAGCGGAGTGGAGTGGAGTGGAGTGGAGTGG 480  
 Db 141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160  
 QY 481 GACCAAGAAGACCAAGTAGATCCGCGCTCATTCATGGGAGATGACCGGCGGAGGAGAC 540  
 Db 161 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 180  
 QY 541 AGCCCTTGGCAGGTGGTCTCTGCTGACTCAAGAAGAAGCTGGCTGGCGGCGGAGTGC 600  
 Db 181 SerProTrpGlnValValLeuLeuLeuAspSerLysLysLysLysLysLysLysLys 200  
 QY 601 ATCCACCCCTCTGGTGTGTGACGCGGCGGCTGATGATGATGATGATGATGATGATGAT 660  
 Db 201 IleHisProSerTrpValLeuThrAlaHisCysMetAspGluSerLysLysLysLys 220  
 QY 661 GTGAGGCTGGAGTATGACCTTGGCGGCTGGGAGAGTGGAGTGGAGTGGAGTGGAGTGG 720  
 Db 221 ValArgLeuGlyGluTrpAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAsp 240  
 QY 721 AAGGAGTCTCTGCTCCACCCCACTACCAAGACGACGACGACGACGACGACGACGACG 780  
 Db 241 LysGluValPheValHisProAsnTrpSerLysSerTrpThrAspAsnAspIleAla 260  
 QY 781 CTGCACTGGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840  
 Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 280  
 QY 841 AGCGGCTTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900  
 Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
 QY 901 GGCTACACGACGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 Db 301 GlyTyrcysSerArgGluLysGluAlaLysArgAsnArgTrpPheValLeuAsnPhe 320  
 QY 961 ATCAAGATTCGCTGGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020  
 Db 321 IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
 QY 1021 GAGAACATGCTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080  
 Db 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
 QY 1081 GGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140  
 Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
 QY 1141 GGTGAGGCTGTGGGCTCTTCAACTAGCGGCGTTCACCAAGTACAGCGGCTACCTTC 1200  
 Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValThrLysValSerArgTrpLeu 400

XX

QY	1	GCCTACTCTCTCTGAGGAGCTCGGTACACAGCAGCTGGAGCGGAGTGCATAGAGAG	60
Db	1	AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu	20
QY	61	ATCTGTGACTTCTGAGGAGCGCCAGGAAATTTTCCAAAATGTGGATGACACACTGGCCTTC	120
Db	21	IleCysAspPheGluGluAlaIysGluIlePheGlnAsnValAspAspThrLeuAlaPhe	40
QY	121	TGGTCCAGCAGCTGCACGGTGACCAAGTCTTGTGTCCTTGGAGCACCCGTGCGCC	180
Db	41	TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla	60
QY	181	AGCCTGTGCTCGGGGCACGGCAGCTGCATCGACGGCATCGGCAGCTTCAGCTCGGACTGC	240



Db 61 SerLeuCysCysGlyHisGlyThrCysIleaspGlyIleGlySerPheSerCysAspCys 80  
 QY 241 CGCAGCGGCTGGAGGCGCTTCGCCAGCGGAGGTGAGCTTCCTCAATTCCTGCTG 300  
 Db 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
 QY 301 GACACGGCGCTGCACGCATTACTGCTAGAGGAGTGGCTGGCGGCGCTGAGCTGT 360  
 Db 101 AspAsnGlyGlyCysThrHisTrpCysLeuGluGluValGlyTrpArgArgCysSerCys 120  
 QY 361 GCGCTGCTCAAGCTGGGAGGACCTCTCTGCAAGTGTACCCCGCAGTGAAGTTCCT 420  
 Db 121 AlaProGlyTrpLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 140  
 QY 421 TGTGGAGCGCTGGAAGCGGATGAGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAG 480  
 Db 141 CysGlyArgProTyrPheArgMetGluLysLysArgSerHisLeuLysArgThrGlu 160  
 QY 481 GACCAAGAGACCAAGTAGATTCGCGCTCATTTGATGGGAAGATGACCGAGGGGAGAC 540  
 Db 161 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgGlyAsp 180  
 QY 541 AGCCCTCTGGCAGGTGGCTGCTGGAGTCAACAGAGAGCTGGCTGGCGGCGAGTCTC 600  
 Db 181 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 200  
 QY 601 ATCCACCCCTCTGGTGGTGTGACAGCGGCGGCTGATGATGATGATGATGATGATGAT 660  
 Db 201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu 220  
 QY 661 GTACAGGCTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 Db 221 ValArgLeuGlyGluTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspLeu 240  
 QY 721 AAGGAGGTCTCTGCTCCACCCCACTACAGCAAGAGCACCACCGCAATGATCGCACTG 780  
 Db 241 LysGluValPheValHisProAsnTrpSerLysSerTrpThrAspAsnAspIleAlaLeu 260  
 QY 781 CTCACCTGGCCAGCGCCGCTCTCGCAGACCATAGTGGCCATCTGCTCCCGGAC 840  
 Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 280  
 QY 841 AGCGGCTTGCAGAGCGGAGTCAATCAGCGCGGCGGAGAGACCTCGTGACGGGTGG 900  
 Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
 QY 901 GGTACCCAGCAGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 Db 301 GlyTyrHisSerArgGluGlyAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
 QY 961 ATCAAGATTCCGCTGGTCCCGCACAATGAGTGCAGCGGAGTGCATGAGCAACATGGTGTCT 1020  
 Db 321 IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
 QY 1021 GAGACATCTGTGTGGGCGCATCTCGGGAGCGGAGGAGTGCCTGCGAGGCGGAGT 1080  
 Db 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
 QY 1081 GGGGGGCGCATGTCGCTCTCTCCAGCGGACCTGCTCTGCTGGGCGCTGTGAGCTGG 1140  
 Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
 QY 1141 GGTAGGCGTGTGGGCTCTTCAACTACGGCGTTTACACCAAGTACAGCGCTACCTC 1200  
 Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTrpLeu 400  
 QY 1201 GACTGGATTCATGGCACATCAGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245  
 Db 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415  
 RESULT 4  
 AAB36894  
 ID AAB36894 standard; Protein; 419 AA.

XX AAB36894;  
 XX 26-FEB-2001 (first entry)  
 DT Human protein C derivative 1.  
 DE  
 DE XX  
 KW Protein C; human; vascular occlusive; burn; transplantation;  
 KW deep vein thrombosis; sickle cell; thalassemia;  
 KW thrombotic disorders; myocardial infarction; angina; stroke.  
 OS Homo sapiens.  
 XX WO200066754-A1.  
 PN 09-NOV-2000.  
 XX 13-APR-2000; 2000WO-US08722.  
 XX 30-APR-1999; 99US-0131801.  
 PA (ELIL ) LILLY & CO ELI.  
 PI Gerlitz BE, Jones BE;  
 DR WPI; 2001-007227/01.  
 DR N-PSDB; AAC83311.  
 XX  
 PT Protein C derivatives, useful for treating vascular occlusive disorder,  
 PT hypercoagulable state, thrombotic disorder and disease states  
 PT predisposing thrombosis, comprises specific amino acid substitutions -  
 PS Claim 1; Page 42-44; 57pp; English.  
 XX  
 CC The present invention relates to a human protein C derivative. The  
 CC protein is useful for treating vascular occlusive disorders.  
 CC hypercoagulable states such as sepsis, disseminated intravascular  
 CC coagulation, purpura fulminans, major trauma, major surgery, burns,  
 CC adult respiratory distress syndrome, transplantation, deep vein  
 CC thrombosis, heparin-induced thrombocytopenia, sickle cell disease,  
 CC thalassemia, viral hemorrhagic fever, thrombotic thrombocytopenic  
 CC purpura, and hemolytic uremic syndrome, and also useful for treating  
 CC thrombotic disorders and acute coronary syndromes such as myocardial  
 CC infarction, unstable angina, and stroke. Protein C derivatives with  
 CC amino acid substitutions result in increased resistance to  
 CC inactivation by serpins when compared to wild-type activated human  
 CC protein C. They also have longer half-lives in human blood and hence  
 CC require either less frequent administration and/or smaller dosage  
 CC than wild type human protein C for treating disorders.  
 XX SQ Sequence 419 AA;

Alignment Scores:  
 Pred. No.: 1.34e-159 Length: 419  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 22 Gaps: 0

US-09-763-153-1 (1-1245) x AAB36894 (1-419)

QY 1 GCCAATCTCTCTGGAGGAGCTCCGTCACAGAGCTGGAGGGAGTGCATAGAGGAG 60  
 Db 1 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 20  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGGAAATTTCCAAATGTGTGATGCACACTGGCCCTC 120  
 Db 21 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspThrLeuAlaPhe 40  
 QY 121 TGTGTCCAAGCAGCTCCAGCGTGACCGTGTGGCTTGGCTTGGCCCTTGGAGCACCCTGCGCC 180  
 Db 41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60

Tue Mar 18 16:19:30 2003

## RESULT 5

AAU99002 standard; Protein; 419 AA.  
 AAU99002; (first entry)  
 23-AUG-2002 (first entry)  
 Human Protein C zymogen protein.  
 Human; Protein C; N-glycosylation; APC; activated protein C; zymogen; serum half-life; chromosome 2q13-q14; stroke; myocardial infarction; after venous thrombosis; disseminated intravascular coagulation; DIC; sepsis; septic shock; embolism; pulmonary embolism; burn; pregnancy; bone marrow transplantation; major surgery; trauma; ARDS; coagulant; adult respiratory distress syndrome; alpha-1 antitrypsin.  
 Homo sapiens.  
 Key Location/Qualifiers  
 Protein 1..155 /label= Light\_chain  
 Peptide 156..157 /label= Lys\_Arg\_dipeptide  
 Peptide 158..169 /label= Activation\_peptide  
 Protein 158..419 /label= Heavy\_chain  
 WO200232461-A2.  
 25-APR-2002.  
 15-OCT-2001; 2001WO-DK00679.  
 18-OCT-2000; 2000DK-0001560.  
 18-OCT-2000; 2000US-242268P.  
 21-JUN-2001; 2001DK-0000970.  
 21-JUN-2001; 2001US-300154P.  
 (MAXY-) MAXYGEN APS.  
 (MAXY-) MAXYGEN HOLDINGS LTD.  
 Andersen KV, Pedersen AH, Freskgaard PO;  
 WPI; 2002-489875/52.  
 N-PSDB; ABK860039.  
 Novel conjugate useful for treating or preventing septic shock, stroke and myocardial infarction, comprises non-polypeptide group covalently attached to protein C polypeptide comprising an attachment group -  
 Claim 2; Page 79-81; 92pp; English.  
 The invention relates to a conjugate (I) comprising at least one non-polypeptide moiety (II) (e.g. an N-glycosyl group) covalently attached to a protein C polypeptide comprising an amino acid sequence which differs from that of a parent protein C polypeptide (III) in at least one introduced group and/or at least one removed amino acid residue comprising an attachment group for the non-polypeptide group (e.g. an N-glycosylation site). Also included are (1) a variant (IV) of (III) comprising a substitution in a position (P) where (P) is an amino acid with at least 25% of its side group exposed to the surface, with the proviso that the substitution is not Thr245Ser/Ala/His/Lys/Arg/Asn/Asp/Glu/Gly/Gln, Tyr302Ser/Ala/Thr/His/Lys/Arg/Asn/Asp/Glu/Gly/Gln or Phe316Ser/Ala/Thr/His/Lys/Arg/Asn/Asp/Glu/Gly/Gln; (2) a nucleotide sequence (V) encoding (IV); (3) an expression vector (VI) comprising (V); (4) a host cell (VII) comprising (V) or (VI); (5) increasing (M2) the functional in vivo half-life of the serum half-life of a parent protein C polypeptide. The conjugates, variants and protein C proteins are useful as medicaments, and in the manufacture of medicaments for the treatment (and diagnosis/prevention) of stroke, myocardial infarction, after venous thrombosis, disseminated intravascular coagulation (DIC), sepsis, septic

181 AGCTGTGCTCGGCGACGGCAGCTGTCATCGAGCGGATCGGCGCTTCAGTCCGACTGC 240  
 61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
 241 CGCAGCGCTGGAGGCGCTTCCTGCGAGCGGAGGTGAGCTTCCTCAATTCCTCGCTG 300  
 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
 301 GACACGGCGGCTGCACCATCTACTGCTAGAGAGGTGGCTGGCGGCTGAGTGT 360  
 101 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 120  
 361 GCGCTGGCTCAAGCTGGGAGGACCTCTCTGAGTGTACCCCGGAGTGAAGTCCCT 420  
 121 AlaProGlyTyrLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 140  
 421 TGTGGAGGCGCTGGAGCGGATGGAGAGAGCGCAGTCACCTGAGACGACACAGAA 480  
 141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160  
 481 GACACAGACCAAGTACATCGGGCTCATTTGATGGAAGATGACCGCGGGGAGAC 540  
 161 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 180  
 541 AGCCCTGGCAGGTGCTCTGCTGGACTCAAGAAGAGTGGCTGGCGGCGGAGTGC 600  
 181 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 200  
 601 ATCCACCCCTCTGGTGTGCTGACAGCGCCCTGATGATGATGATGATGATGATGATGAT 660  
 201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 220  
 661 GTCAGGCTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 221 ValArgLeuGlyGlyTyrAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 240  
 721 AAGAGGTCTGCTGCGCCCACTACAGCAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 780  
 241 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 260  
 781 CTGACCTGCGCCGCGCGCCGCGCCGCGCCGCGCCGCGCCGCGCCGCGCCGCGCCGCG 840  
 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 280  
 841 AGCGGCTTGGAGCGGCGGCTCAATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
 901 GGCTACACAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 301 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
 961 ATCAAGATTCGCTGGTCCCGCACAATCAGTGTGCGAGGCTGATGACCAATGCTGTCT 1020  
 321 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
 1021 GAGAACATCTGTGCGGCGATCTCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
 1081 GGGGGGCGCATGCTGCTCTCTTCCACGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 361 GlyClyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
 1141 GGTGAGGCTGTGGTCTCTCTCAACTACGCGCTTACACAAAGTCAGCGGCTACCTC 1200  
 381 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 400  
 1201 GACTGGATCATGGGACATCATGAGACAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245  
 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415

CC shock, emboli e.g. pulmonary emboli, transplantation such as bone marrow  
 CC transplantation, burns, pregnancy, major surgery/trauma or adult  
 CC respiratory distress syndrome (ARDS). The variant protein C has an  
 CC increased resistance to activation by e.g. human plasma and alpha-1  
 CC antitrypsin. The conjugates have an increased in vivo half-life,  
 CC increased serum half-life, increased resistance to inhibitors, reduced  
 CC renal clearance, reduced immunogenicity and/or increased bioavailability.  
 CC The conjugate offers a number of advantages over the currently available  
 CC APC products, including longer duration between injections,  
 CC administration of less protein, and fewer side effects. Moreover, a  
 CC reduced anticoagulant activity is beneficial to reduce the risk of  
 CC bleeding while maintaining the antiinflammatory activity of APC  
 CC (activated protein C) conjugates. This must be especially important when  
 CC the conjugate has an extended plasma life. The gene for protein C is  
 CC located on chromosome 2q13-q14. The present sequence represents zymogen  
 CC protein C upon which the variants of the invention were based.

Sequence 419 AA;

# Alignment Scores:

Pred. No.: 1.34e-159 Length: 419  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 23 Gaps: 0

US-09-763-153-1 (1-1245) x AAU99002 (1-419)

QY	1	GCACACTCTCTCCTGAGAGCTCGTCCACAGCAGCTGGAGCGGAGTGCATAGAGGAG	60
Db	1	AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu	20
QY	61	ATCTGTGACTTGGAGAGGCCAAGAAATTTTCCAAATFTGGATGACACACTGGCCTTC	120
Db	21	IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe	40
QY	121	TGGTCCAAAGCAGCTGAGGTTGACAGTGTGCTTGTCTGCTTGGACACCGCTGCGCC	180
Db	41	TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla	60
QY	181	AGCTGTGCTCGGGCAGCGACGTGTCATCGAGCGATCGGACCTTCAGCTCGGACTGC	240
Db	61	SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys	80
QY	241	CGCAGCGCTGGAGGCGGCTTCCACCGCGAGGTGAGCTTCCTCAATTGCTGCTG	300
Db	81	ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu	100
QY	301	GACAAACGGCGCTGCAGCATTACTGCTAGAGAGGTGGCTGGCGGCTGTAGCTGT	360
Db	101	AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys	120
QY	361	CGCGCTGGCTACAGCTGGGGGACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	420
Db	121	AlaProGlyTyrLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro	140
QY	421	TGTGGAGGCCCTTGAAGCGGATGGAGAAAGACGACGTCACCTCAACAGCAGACAGAA	480
Db	141	CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu	160
QY	481	GACCAAGAAGCAGTAGATCCGGCTGCTATGATGGAAGATGACCGAGCGGGGAGAC	540
Db	161	AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp	180
QY	541	AGCCCTGGCAGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	600
Db	181	SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu	200
QY	601	ATCCACCCCTCTGGTGTGACAGCGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	660
Db	201	IleHisProSerTrpValLeuThAlaAlaHisCysMetAspGluSerLysLysLeuLeu	220

QY	661	GTCAGGCTTGAGAGATGATGACCTCGCGCTGGGAGAGTGGAGCTGGACCTGGACATC	720
Db	221	ValArgLeuGlyGluTyrAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle	240
QY	721	AAGAGGCTCTCTGCTCCACCCCACTACGACAGAGCACCACCAATGACATCGCACCTG	780
Db	241	LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu	260
QY	781	CTGCACCTGGCCAGCCCGCCCTCTGCGACACCATAGTCCCATCTGCTCCCGGAC	840
Db	261	LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp	280
QY	841	AGCGGCTTTCAGAGCGGAGCTCAATCAGGCGCGGAGAGACCTCTGTCAGCGCTGG	900
Db	281	SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp	300
QY	901	GGCTACACAGCAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	960
Db	301	GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe	320
QY	961	ATCAAGATTCCGCTGCTCCCGCACATGAGTGCAGCGAGGTGCATGAGCAACATGCTGT	1020
Db	321	IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer	340
QY	1021	GAGACATGCTGTGCGGCGATCTCTCGGGACCGGAGGAGGAGGAGGAGGAGGAGGAG	1080
Db	341	GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer	360
QY	1081	GGGGGCGCCATGGTCCCTCTTCCACGGCACCTGGTTCCTGCTGGCTGGCTGGCTGG	1140
Db	361	GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp	380
QY	1141	GCTGAGGCTGTGGGCTCTCTCAACTACGCGCTTTACACCAAGTACCGCTACCTC	1200
Db	381	GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerA-gTyrLeu	400
QY	1201	GACTGATTCATGGGACATCAGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1245
Db	401	AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415	

## RESULT 6

AAP81104  
 ID AAP81104 standard; protein; 460 AA.

AC AAP81104;

XX 16-SEP-1990 (first entry)

DE Sequence of human protein C.

XX Human protein C; plasmid ppc 1.

OS Homo sapiens.

XX JP63263083-A.

XX 30-NOV-1988.

PD 21-APR-1987; 87JP-0096341.

PF 21-APR-1987; 87JP-0096341.

PR (FARH ) HOECHST JAPAN KK.

XX WPI; 1988-350711/49.

DR N-PSDB; AAN81408.

XX Human protein C gene - prepd. from new DNA of specified base sequence.

PS Disclosure; ; 16pp; japanese.

XX The human protein C is expressed in large amts. using plasmid ppc 1 in

CC E.coli K12/Om 225 (FERM P-9297).

US-09-763-153-1 (1-1245) x AAP81104 (1-460)

[illegible]

CC using the DNA construct to breed a transgenic animal (esp. sheep,  
 CC rabbit, cattle, goat) that produces protein C in its milk, at  
 CC least 90% of the protein C being in the two-chain form.  
 CC Modification of the protein C two-chain cleavage site (see AAW25085)  
 CC improves the maturation of recombinant protein C from single chain  
 CC to two-chain form.  
 XX SQ Sequence 460 AA;

Alignment Scores:

Pred. No.: 1.37e-159 Length: 460  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 Gaps: 0

S-09-763-153-1 (1-1245) x AAW25086 (1-460)

QY	1	GCACACTCTCTCTGGAGGAGCTCCGTCACAGCAGCTGGAGCGGAGTGCATAGAGAG	60
Db	43	AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGlu	62
QY	61	ATCTGTCACTTCAGAGAGCCAGGAATTTTCCAAATGTGGATGACACACTGGCCTTC	120
Db	63	IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe	82
QY	121	TGTTCAAGCAGCTGACGAGTGCACGAGTGTGGTCTGGTCTGGTGGAGCCCGTGGCC	180
Db	83	TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla	102
QY	181	AGCTGTGTGGGCGGCGGAGTGCATGTCAGCGCATCGGAGCTTCAGCTGCAGTGC	240
Db	103	SerLeuCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys	122
QY	241	CGCAGCGCTGGAGGCGGCTTCTCCAGCGGAGTGCATGTCAGCGCATCGGAGCTGC	300
Db	123	ArgSerGlyTrpGluArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu	142
QY	301	GACACGGGCGGCTGACCACTTACTGCTAGAGGAGTGGGCTGGGCGGCTGAGCTGT	360
Db	143	AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgGlySerCys	162
QY	361	GGCCTGTGCTACAGCTGGGCGGAGTGCATGTCAGCGCATCGGAGCTTCAGCTGC	420
Db	163	AlaProGlyTyrLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro	182
QY	421	TCTGGGAGGCGCTGAGAGCGGATGGAGAGCGGAGTGCATGTCAGCGCATCGGAG	480
Db	183	CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu	202
QY	481	GACCAAGAGACCAAGTAGATCCGCGCTCATGTGGAAGATGACAGCGGCGGAGAC	540
Db	203	AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp	222
QY	541	AGCCCTGGCAGGTGCTGCTGGACTCAAGAGAGAGTGGCTGGGCGGAGTGCCTC	600
Db	223	SerProTrpGlnValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu	242
QY	601	ATCCACCCCTCTCTGGTGTGCTGACAGCGGCGCTGAGTGCATGATGATCAAGAGCTCT	660
Db	243	IleHisProSerTrpValLeuThrAlaLaHisCysMetAspGluSerLysLysLeuLeu	262
QY	661	GTACAGCTGGAGATGATGACCTGGGCGGCTGGGAGAGTGGAGCTGGACCTGGACATC	720
Db	263	ValArgLeuGlyGluTrpAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle	282
QY	721	AAGGAGTCTTCTGCTCCACCCCACTACAGAGAGCACCACCCAGATGACATCGCACTG	780
Db	283	LysGluValPheValHisProAsnTyrSerLysSerThrAspAsnAspIleAlaLeu	302
QY	781	CTGCACCTGGCCCGGCGGAGCTTCGCGAGACCATAGTCCCATCTGCTCCGCGAC	840

Db	303	LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp	322
QY	841	AGCGGCTTGCAGAGCGGAGTCAATCAGCGCGGAGGAGCCCTGCTGACGGGCTGG	900
Db	323	SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp	342
QY	901	GGTACCACAGCAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	960
Db	343	GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe	362
QY	961	ATCAAGATTCCTGGTGGTCCCGCACAAATAGTGCAGGAGGTGATGAGCAACATGGTCT	1020
Db	363	IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer	382
QY	1021	GAGAACATGCTGTGTGGGCGATCTCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAG	1080
Db	383	GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer	402
QY	1081	GGGGGCGGCGGCTTCTCCAGCGGAGTGCATGTCAGCGCATCGGAGCTTCAGCTGC	1140
Db	403	GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp	422
QY	1141	GGTAGGCTGTGGGCTCTTCACAACTAGCGGCTTACACCAAGTACAGCGCTACCTC	1200
Db	423	GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValThrLysValSerArgTyrLeu	442
QY	1201	GACTGGATCCATGGGCGACATCAGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1245
Db	443	AspIleHisGlyHisIleArgAspLysGluAlaProGlnLys	457

RESULT 8

AAP60001  
 ID AAP60001 standard; Protein; 461 AA.

AC AAP60001;

XX 25-JUL-1991 (first entry)

DE Sequence of polypeptide with human protein C activity.

XX Vascular disorder therapy; protein C deficiency.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Region 1..32

FT Protein /note="encoded by AAN60004"

FT /note="33..461"

XX /note="encoded by AAN60001"

PN EP191606-A.

XX 20-AUG-1986.

XX 06-FEB-1986; 86EP-0300823.

XX 08-FEB-1985; 85US-0699967.

XX (EIL) ELI LILLY & CO.

XX Bang NU, Beckmann RJ, Jaskunas SR, Lai MHT, Little SP;

PI Long GL, Santerre RF;

XX WPI; 1986-220077/34.

XX Prodn. of polypeptide having human protein C activity - is by

XX recombinant DNA procedures for prod. useful against vascular

XX disorders

PS Disclosure; Pages 10-12; 121pp; English.

XX The claimed sequence AAN60001 has "RIN-RW" attached to its 5' end

CC wherein: R= AAN60002 or AAN60003, and RI= AAN60004 or AAN60005; and M and

Tue Mar 18 16:19:30 2003

CC N= 0 or 1; provided that when M=0, N=0; and that when R= AAN60002, R1= AAN60004; and that when R= AAN60003, R1= AAN60005.

XX SQ Sequence 461 AA;

Alignment Scores:  
Pred. No.: 1.37e-159 Length: 461  
Score: 2298.00 Matches: 415  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.33% Indels: 0  
DB: Gaps: 0

US-09-763-153-1 (1-1245) x AAP60001 (1-461)

QY 1 GCCAACTCTCTCTCGAGAGCTCGCTCACAGACCTCGAGCGGAGTGCATAGAGGAG 60  
DB 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
QY 61 ATCTGTGACTTCGAGAGGCAAGAAATTTCCAAAATGTGGATGACACACTGGCTTC 120  
DB 63 IleCysAspPheGluGluAlaLysGluLeuPheGlnAsnValAspThrLeuAlaPhe 82  
QY 121 TGTCTCAAGCAGCTGACGCTGACAGTCTGTGTCTTGCCTTGGAGCACCCTGGCC 180  
DB 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
QY 181 AGCTGTGCTCGCGGACGCGACGCTCATCGACGCGCATCGGAGCTTCAGCTGCGACTGC 240  
DB 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
QY 241 CGCAGCGCTGGAGGCGCTTCTGCCAGCGGAGGTGAGTCTCTCAATGTCTCGCTG 300  
DB 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
QY 301 GACACGCGCTCCAGCGCTTACTGCTAGAGAGGTGGCTGGCGGCGCTGTAGCTGT 360  
DB 143 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 162  
QY 361 GCGCTGCTCAAGCTGGGAGGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 163 AlaProGlyTyrLysLeuGlyAspLeuGlnCysHisProAlaValLysPhePro 182  
QY 421 TGTGGAGGCTTGAAGCGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
DB 183 CysGlyArgProTrpPlyArgGluGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
QY 481 GACCAAGAGACCAAGTATGATCCGCGCTCATGATGGGAAGATGACCGGCGGAGAC 540  
DB 203 AspGlnGluAspGlnValAspProArgLeuLeuAspGlyLysMetThrArgArgGlyAsp 222  
QY 541 AGCCCTGCGAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
DB 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242  
QY 601 ATCCACCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
DB 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 262  
QY 661 GTACGCTTGGAGCTATGACCTGCGGCGCTGGGAGAGTGGAGCTGGAGCTGGACATC 720  
DB 263 ValArgLeuGlyGlyTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
QY 721 AAGAGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
DB 283 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 302  
QY 781 CTGACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840  
DB 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
QY 841 AGCGGCTTGCAGAGCGGAGCTCAATCAGCGCGGCGGCGGCGGCGGCGGCGGCGG 900

DB 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
QY 901 GGCTACACAGAGCCGAG 960  
DB 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
QY 961 ATCAGATTCCCGTGTCCCGCACAAATGAGTGCAGCGAGGTTCATGACACATGGTGTCT 1020  
DB 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
QY 1021 GAGACATGCTGTGTGCGGCGATCTCCGGGACCGCGGAGGATGCTCGGAGGCGACAGT 1080  
DB 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
QY 1081 GGGGGCCCATGTGCTGCTCTCCACGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
DB 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
QY 1141 GGTGAGGCTGTGGGTCTCTTCAACTAGCGCTTACACCAAGTACAGCGCTACCTC 1200  
DB 423 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 442  
QY 1201 GACTGGATCCATGGCACATCAGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245  
DB 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457  
RESULT 9  
AAP70855  
ID AAP70855 standard; Protein: 461 AA.  
XX AC AAP70855;  
XX 10-MAY-1991 (first entry)  
DT Human Protein C.  
XX human Protein C; anti-coagulant; thrombosis; serine protease.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FT Peptide 1..42 /label= prepro leader peptide  
FT Disulfide-bond 59..64 /label= gamma-carboxyglutamic acid (Gla) domain  
FT Domain 60..63 /label= growth factor domains  
FT Disulfide-bond 92..111 /label= growth factor domains  
FT Disulfide-bond 101..106 /label= growth factor domains  
FT Disulfide-bond 105..120 /label= growth factor domains  
FT Disulfide-bond 122..131 /label= growth factor domains  
FT Disulfide-bond 140..151 /label= growth factor domains  
FT Disulfide-bond 147..160 /label= growth factor domains  
FT Disulfide-bond 162..175 /label= growth factor domains  
FT Domain 92..175 /label= growth factor domains  
FT Disulfide-bond 183..319 /label= growth factor domains  
FT Disulfide-bond 373..387 /label= growth factor domains  
FT Disulfide-bond 398..426 /label= growth factor domains  
FT Disulfide-bond 238..254 /label= growth factor domains  
FT Modified-site 139 /label= N-glycosylation site  
FT Modified-site 290 /label= N-glycosylation site  
FT Modified-site 355 /label= N-glycosylation site  
FT Modified-site 371 /label= N-glycosylation site  
FT Modified-site 211..212 /label= N-glycosylation site  
FT Cleavage-site /note= "in heavy chain; converts to activated protein C"  
FT Cleavage-site 197..198 /note= "apparent processing site for connecting

FT Cleavage-site 199..200 dipeptide "  
 FT /note= "apparent processing site for connecting  
 FT dipeptide"  
 XX EP215548-A.  
 XX 25-MAR-1987.  
 XX 26-JUN-1986; 86EP-0304970.  
 XX 27-JUN-1985; 85US-0749600.  
 PR 15-AUG-1985; 85US-0766109.  
 XX (ZYMO-) ZYMOGENETICS INC.  
 PA (UNIW ) UNIV OF WASHINGTON.  
 XX Murray MJ, Berkner KL, Foster DC, Davie EW;  
 DR WPI; 1987-081505/12.  
 DR N-PSDB; AAN70102.  
 XX  
 PT Human protein C or activated protein C - prepd. using expression  
 PT vector capable of integration in mammalian host cell DNA  
 PS Claim 4; Fig 4; 52pp; English.  
 XX  
 CC Recombinantly produced protein C can be used to treat thrombotic  
 CC disorders such as venous thrombosis as it has anti-coagulant  
 CC properties. The protein sequence is thought to yield two peptide  
 CC chains; the first contains the Gla domain and growth factor domains  
 CC and the second (the activation peptide) contains the catalytic  
 CC domain.  
 XX  
 SQ Sequence 461 AA;

Alignment Scores:  
 Pred. No.: 1,37e-159 Length: 461  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservativity: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 8 Gaps: 0

US-09-763-153-1 (1-1245) x AAP70855 (1-461)

1 GCCAACTCTTCCTGGAGAGCTCGTCACACAGCCCTGGAGCGGAGTGCATAGAGGAG 60  
 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluA-gGluCysIleGluGlu 62  
 61 ATCTGTGACTTCGAGGAGGCCAAGGAAATTTCCAAATGTCGATGACACACTGGCCTTC 120  
 63 IleCysAspPheGluGluAlaAlaGluLeuPheGlnAsnValAspAspThrLeuAlaPhe 82  
 121 TGGTCCAAAGCAGCTGACCGTGACAGTGTGCTTGTGCTTGGAGCAGCCGTCGCC 180  
 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
 181 AGCTGTGCTGGCGCAGCGACGTCGATCGACGGCATCGGAGTTCAGCTTCGACTGC 240  
 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
 241 CGCAGCGCTGGAGGCGCGCTTCTGCCAGCGGAGGTGAGCTTCTCAATTGCTCGCTG 300  
 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
 301 GACAAGCGGCTGCACGCATTAAGTCTAGAGAGTGGGCTGGCGGCTGTAGCTGT 360  
 143 AsnAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgCysSerCys 162  
 361 CGCCTGGCTACAAGCTGGGGGACGACCTCTGCTGCTGTCACCCCGCTGAGTTCCT 420  
 163 AlaProGlyTyrLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro 182

QY 421 TGTGGAGAGCCCTGGAGCGGATGGAGAGAGAGCGCAGTCACTGAAACGAGACACAGAA 480  
 Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
 QY 481 GACCAAGAAGACCAAGTAGATCCGCGCTCATTTGATGGGAAGATGACACAGCGGGAGAC 540  
 Db 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
 QY 541 AGCCCTGGCAGGTGCTCTGCTGACCTCAAGAAGAGTGGCTGCGGGGAGCTGCTC 600  
 Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 242  
 QY 601 ATCCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 660  
 Db 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLeuLeu 262  
 QY 661 GTCAGGCTTGGAGATGATGACCTGCGCGCTGGGAGAGTGGGAGCTGGACCTGGACATC 720  
 Db 263 ValArgLeuGlyGluTyrAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
 QY 721 AAGGAGTCTTCGCTCCACCCCACTACAGCAAGAGCACCACCAATGACATCGCAGCTG 780  
 Db 283 LysGluValPheValHisProAsnTyrSerLysSerThrAspAsnAspIleAlaLeu 302  
 QY 781 CTGACCTGGCCCGCCAGCCGCTCTCCGACACCATAGTGCCTCTGCTCCCGGAC 840  
 Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
 QY 841 AGCGGCTTCGAGAGCGGAGCTCAATCAGCGCGGCGGAGAGACCTCTGACGGCTGG 900  
 Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
 QY 901 GGCTACACAGCAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 Db 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
 QY 961 ATCAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
 Db 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
 QY 1021 GAGAACATGCTGTGCGGGGATCTCTCGGGGACCGGAGGAGTGCCTGCGAGGCGGACAGT 1080  
 Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
 QY 1081 GGGGGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
 QY 1141 GGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
 Db 423 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 442  
 QY 1201 GACTGATCCATGGGCACATCAGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245  
 Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

RESULT 10  
 AAP90401  
 ID AAP90401 standard; protein; 461 AA.  
 XX  
 AC AAP90401;  
 XX  
 DT 01-NOV-1989 (first entry)  
 XX  
 DE Zymogen form of human protein C.  
 XX  
 KW Human protein C; zymogen form; activated C protein;  
 XX human liver mRNA; signal peptide; propeptide; antithrombotic.  
 OS Homo sapiens.  
 XX  
 PN EP323149-A.

241	QY	CGCAGCGCTGGAGGCGCGCTTCTGCCAGCGCGAGGTGAGCTTCCTCAATTGCTCGCTG	300
122	Db		
123	Db	ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu	142
301	QY	GACAAACGCGCGCTGCACGATTAATCTAGAGGAGGTGGCTGGCGCGCTCTAGCTGCT	360
143	Db		
361	QY	CGGCTGGCTACAGCTGGGGAGACACCTCTCTGCAGTGTCAACCCCGAGTGAAGTTCCCT	420
163	Db		
421	QY	TGTGGAGGCGCCTGGAAAGCGGATGGAGAGACGCGACCTCACTGAAACGAGACACAGAA	480
183	Db	CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgPheThrGlu	202
481	QY	GACCAAGAGACCAAGTAGATCCGCGCTCATTTGATGGGAAGATGACACAGCGGGAGAC	540
203	Db	AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp	222
541	QY	AGCCCTGGCAGGTGCTCTGCTGGACTCAAGAGAGAGCTGGCTGGCGGGCAGTGCCTC	600
223	Db	SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu	242
601	QY	ATCCACCCCTCTGGTGCTGACAGCGGCGCCACTGCGATGGATGATCCAAAGACTCCTT	660
243	Db	IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu	262
661	QY	GTGAGCTTGGAGATGACTCGCGCGCTGGGAGAAGTGGAGCTGGACTCGACTCGACATC	720
263	Db	ValArgLeuGlyLutryAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle	282
721	QY	AAGGAGGTCTTCGTCCACCCCACTACAGCAAGACGACACCGCAATGACATCGCACTG	780
283	Db	LysGluValPheValHisProAsnTrpSerLysSerThrThrAspAsnAspIleAlaLeu	302
781	QY	CTGCACTGGCCCCAGCCGCGCCTCTCGCAGACCATAGTGCCTATCTGCCTCCCGAC	840
303	Db	LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp	322
841	QY	AGCGGCTTGCAGAGCGGAGCTCAATCAGCGCGCGCAGGAGACCTCTGACGGGCTGG	900
323	Db	SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp	342
901	QY	GGCTACCACAGCAGCGAGAGAGGAGGCGCAAGAGAAACCGCACCTTCGCTCCTCAACTTC	960
343	Db	GlyTrpHisSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe	362
961	QY	ATCAAGATTCCGTGGTCCCGCACATAGTGCAGCGAGTCAATGACCAACATGGTGCTCT	1020
363	Db	IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer	382
1021	QY	GAGAACATGCTGTGTCGGGCACTCTCGGGGACCGGACGAGTGCCTCGGAGCGGACAGT	1080
383	Db	GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer	402
1081	QY	GGGGGCGCATGGTGCCTCTTCCACGGCAGCTGGTTCCTGGTGGCGCTGGTGAAGCTG	1140
403	Db	GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp	422
1141	QY	GGTAGGCTGTGGCTCTTCACTACGGCTTTACCAAGTAGTCAGCGCTACCTC	1200
423	Db	GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValThrLysValSerArgTyrLeu	442
1201	QY	GACTGGATCCATGGGCATTCAGAGACAGGAAGCCCCCAGAAG	1245
443	Db	AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys	457
RESULT 11			
AAAL13074			
ID AAAL13074 standard; Protein: 461 AA.			
XX			



AC AAR13074;  
 XX 02-OCT-1991 (first entry)  
 XX Protein C precursor.  
 XX Anticoagulant; fibrinolysis.  
 XX Homo sapiens.  
 XX Key  
 XX Location/Qualifiers  
 XX Peptide  
 XX 2..42  
 XX /label= pre-pro peptide  
 XX Region  
 XX 43..197  
 XX /label= light chain  
 XX Cleavage-site  
 XX 197..198  
 XX /label= proteolytic cleavage  
 XX Cleavage-site  
 XX 199..200  
 XX /label= proteolytic cleavage  
 XX Peptide  
 XX 201..211  
 XX /label= activation peptide  
 XX Region  
 XX 212..461  
 XX /label= heavy chain  
 XX Domain  
 XX 43..79  
 XX /label= Gla domain  
 XX Modified-site  
 XX 48  
 XX /label= gamma carboxyglutamic acid  
 XX Modified-site  
 XX 49  
 XX /label= gamma carboxyglutamic acid  
 XX Modified-site  
 XX 56  
 XX /label= gamma carboxyglutamic acid  
 XX Modified-site  
 XX 58  
 XX /label= gamma carboxyglutamic acid  
 XX Modified-site  
 XX 61  
 XX /label= gamma carboxyglutamic acid  
 XX Modified-site  
 XX 62  
 XX /label= gamma carboxyglutamic acid  
 XX Modified-site  
 XX 67  
 XX /label= gamma carboxyglutamic acid  
 XX Modified-site  
 XX 68  
 XX /label= gamma carboxyglutamic acid  
 XX Modified-site  
 XX 71  
 XX /label= gamma carboxyglutamic acid  
 XX Modified-site  
 XX 139  
 XX /label= N-glycosylation site  
 XX Modified-site  
 XX 250  
 XX /label= N-glycosylation site  
 XX Modified-site  
 XX 355  
 XX /label= N-glycosylation site  
 XX Modified-site  
 XX 371  
 XX /label= N-glycosylation site  
 XX W09109951-A.  
 XX 11-JUL-1991.  
 XX 21-DEC-1990; 90WO-US07617.  
 XX 22-DEC-1989; 89US-0456092.  
 XX (TEIJ) TEIJIN LTD.  
 XX (ZYMO-) ZYMOGENETICS INC.  
 XX Foster DC, Holly RD, Suzuki M, Wakabayashi K, Kumar AA;  
 XX WPI; 1991-222903/30.  
 XX N-PSDB; AAQ12649.  
 XX Recombinant protein C with truncated light chain - for use as an  
 XX anticoagulant.  
 XX Disclosure; Fig 1; 60pp; English.

CC The sequence was deduced from a clone isolated from a cDNA library  
 CC prep'd. from mRNA from Hep G2 cells. It is a protein C precursor,  
 CC including light and heavy chains, which is cleaved to produce  
 CC activated protein C (see feature table). The DNA encoding the  
 CC sequence can be manipulated by genetic engineering techniques to  
 CC express a protein comprising (when activated) a heavy chain and a  
 CC truncated light chain comprising residues 1-149, 1-150, 1-151 or 1-  
 CC 152 of the natural sequence. The protein pref. comprises the  
 CC precursor of formula:  
 CC  
 CC pre-pro = pre-pro-L-X-H  
 CC The corresponding peptide of protein C with all/part replaced by  
 CC or prothrombin;  
 CC L = Abs 1-149, 150, 151 or 152 of light chain;  
 CC X = 3-10 Lys/arg residues; and  
 CC H = heavy chain.  
 CC Cells transformed with expression vectors contg. the modified DNA  
 CC sequences produce the new proteins which can be used to regulate  
 CC anticoagulant and fibrinolytic systems.  
 CC See also W09112320 (AAR13074).  
 XX  
 XX Sequence 461 AA;  
 SQ  
 Alignment Scores:  
 Pred. No.: 1.37e-159 Length: 461  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 12 Gaps: 0  
 US-09-763-153-1 (1-1245) x AAR13074 (1-461)  
 QY 1 GCCAACTCTCTCTGGAGGAGCTCCGTACAGCAGCTGGAGGGAGTGCATAGAGGAG 60  
 Db |||||  
 QY 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
 Db |||||  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGATGATGATGATGATGATGAT 120  
 Db |||||  
 QY 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
 Db |||||  
 QY 121 TGGTCCAGCAGCTCGAGCGTGACCGTGTGCTTGTCTTGCCTTGGAGCACCCTGGCCG 180  
 Db |||||  
 QY 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
 Db |||||  
 QY 181 AGCCTGTCTCGGGGACGCGCAGTGCATCGAGCGGATCGGCGAGTTCAGTGGAGTGC 240  
 Db |||||  
 QY 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
 Db |||||  
 QY 241 CGCAGCGCTGGGAGGGCGCTTCTCCAGCGAGGAGTTCCTCAATTGCTCGCTG 300  
 Db |||||  
 QY 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
 Db |||||  
 QY 301 GACACGGCGGCTGACGACGATTCCTCTAGAGAGGAGTGGCGTGGCGCGCTGTAGCTGT 360  
 Db |||||  
 QY 143 AspAsnGlyGlyCysThrHisThrCysLeuGluValGlyTrpArgArgCysSerCys 162  
 Db |||||  
 QY 361 GCGCCTGGCTACAGCTGGGGGACGACCTCTGAGTGTACCCCGAGTGAAGTTCCT 420  
 Db |||||  
 QY 163 AlaProGlyTrpLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro 182  
 Db |||||  
 QY 421 TGTGGAGGCCCTTGAAGCGGATGGAGAGAGCGGAGTACCTGAGAACGAGACAGAA 480  
 Db |||||  
 QY 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
 Db |||||  
 QY 481 GACCAAGAGACCAAGTAGATCCCGGCTCATTTGATGGAGATGACACAGCGGGGAGAC 540  
 Db |||||  
 QY 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
 Db |||||  
 QY 541 AGCCCTTGGCAGGTGGCTCTGCTGACTCAAGNAGAGCTGCCTGCGGGGAGTGTCTC 600  
 Db |||||  
 QY 223 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 242  
 Db |||||

601	ATCCACCCCTCCCTGGTGTGTGACAGCGGCCACTGCATGGATGAGTCCAGAGAGCTCCTT	660
243	IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu	262
661	GTCAAGGCTTGGAGATGACCTCGCGGCTGGGAGAAGTGGAGCTGGACCTCGACATC	720
263	ValArgLeuGlyGluTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle	282
721	AAGAGGCTTCGTCCACCCCACTACAGACAGACACCGACCAATGACATCGCACTG	780
283	LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu	302
781	CTGCACCTGGCCCGCGCCACCTCTCCACACCATAGTCCCATCTGCCTCCCGGAC	840
303	LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp	322
841	AGCGGCTTGGACGCGGAGCTCAATCAGCGCGCCAGAGACCCCTCGTACGGGCTGG	900
323	SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp	342
901	GGCTACCCAGCAGCGGAGAGAGAGGCGCCAAAGAGAACCACCTTCGTCTCAACTTC	960
343	GlyTyrHisSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe	362
961	ATCAAGATCCCGTCCCGCACAATGAGTGCAGCGAGGTGCATGAGCAACATGGTGCT	1020
363	IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer	382
1021	GAGNACATGCTGTGTCGGGCATCTCTCGGGGACCGCAGGATGCTCGGAGGCGACAGT	1080
383	GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer	402
1081	GGGGGGCCCATGTGCGCTCTTCACACTACGGCACCCTGGTTCCTGGTGGCCTGGT	1140
403	GlyGlyProMetValAlaSerPheHisGlyThrThrPheLeuValGlyLeuValSerTrp	422
1141	GCTCAGGCTGTGGCTTCCTTCAACTACGGCGCTTTACACCAAGTACAGCGGTACTCT	1200
423	GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu	442
1201	CAGTGTGATCCATGGGCACATCAGACACAGGAGCGCCCGCAGAAG	1245
443	AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys	457
RESULT 12		
AAIR13081	AAIR13081 standard; Protein; 461 AA.	
XX	AC	AAIR13081;
XX	DT	30-SEP-1991 (first entry)
XX	XX	Human protein C.
XX	XX	Phospholipid; binding protein; lipocortin; domain; vitamin K; PBP;
XX	XX	gla-domain; VKDP.
XX	OS	Homo sapiens.
XX	XX	Location/Qualifiers
XX	XX	1..42
XX	XX	/label= sig_peptide
XX	XX	43..461
XX	XX	/label= mat_protein
XX	XX	W09105953-A.
XX	XX	11-JUL-1991.
XX	XX	13-DEC-1990; 90WO-US07335.
XX	XX	29-DEC-1989; 89US-0459082.
XX	XX	

(ZYMO-) ZYMOGENETICS INC.

Foster DC;

WPI; 1991-222905/30.  
N-PSDB; AAQ12678.

Recombinant prodn. of hybrid phospholipid-binding proteins -  
comprising lipocortin phospholipid-binding domain and  
vitamin K-dependent protein

Disclosure; Fig 2; 57pp; English.

This sequence, or a fragment of it, is used in the construction of  
hybrid phospholipid-binding proteins (PBP) having the same biological  
activity as human protein C or human activated protein C.  
The hybrid sequence would comprise at least one lipocortin phospholipid  
binding domain (PBD), e.g. of PAP-I, joined to a gla-domainless  
protein C or activated protein C. See AAQ12680-81 for such examples.  
See also AAQ12678-81.

Sequence 461 AA:

Alignment Scores:  
Pred. No.: 1.37e-159 Length: 461  
Score: 2298.00 Matches: 415  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.33% Indels: 0  
Gaps: 0  
DB: 12

US-09-763-153-1 (1-1245) x AAR13081 (1-461)

QY 1 GCCAACTCCTTCGTGAGGAGCTCGTCCACAGCAGCTGGAGGGGAGTGATAGAGGAG 60  
Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
QY 61 ATCTGTCACTTCGAGGAGGCCAAGGAATAATTTCAAAATGTGGATGACACACTGGCCTTC 120  
Db 63 IleCysaspPheGluGluAlaLysGluIlePheGlnAsnValaspThrLeuAlaPhe 82  
QY 121 TGGTCCAAGCACGTCGACGCTGCACAGTCTGTGGTTCTTGGCCCTTGGAGCACCCGTGGCC 180  
Db 83 TrpSerLysHisValaspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
QY 181 AGCCFTGTCTCGGGGACGGCAGCTGCATCGACGGCATCGGCAGCTTCAGCTCGGACTGC 240  
Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
QY 241 CCACAGCGCTGGAGGGCCGCTTCGCCAGCGGAGGTGAGCTTCTCAATTGCTCGCTG 300  
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
QY 301 GACACGGCGGCTGCACSCATTACTGCCTAGAGAGAGGTGGCTGGCGGCTGTAGCTGT 360  
Db 143 AspAsnGlyGlyCysThrHisTyrcysLeuGluGluValGlyTrpArgargcysSerCys 162  
QY 361 GCSCCTGGCTACAAGCTGGGGAGCACTCTCGAGTGTCCACC CGCAGTGAAGTTCCT 420  
Db 163 AlaProGlyTyrllysLeuGlyAspAspleuleuGlnCysHisProhlaVallysPhepro 182  
QY 421 TGTTGGAGGCCCTTGAAGCGGATGAGAAGAGCGCAGTCCACTGAAACGACACACAGAA 480  
Db 183 CysGlyArgProtrpIysargMetGluylsLysargSerHisLeulysargaspThrGlu 202  
QY 481 GACCACAAGACCAAGTAGATCCGCGGTCAATTGATGGGAAGATGACACAGCGGGAGAC 540  
Db 203 AspGlnGluaspGlnValaspProargLeulleaspGlyIysmethrargargGlyasp 222  
QY 541 AGCCCCGTGCAGTGTCTGCTGCTGACTCAAAGAAGAAGTGGCTCGCGGGCGAGTGCTC 600  
Db 223 SerProtrpGlnValVallleuLeuaspSerLysLysLysLeuAlaCysGlyAlaValLeu 242



```

/label= C-terminal
/note= "light chain"
200..211
/label= N-terminal
/note= "heavy chain"
458..461
/label= C-terminal
/note= "heavy chain"
451..461
/label= C-terminal
/note= "heavy chain"

```

JP05064588-A.  
19-MAR-1993.  
14-AUG-1991; 91JP-0228687.  
14-AUG-1991; 91JP-0228687.  
(TEIJ ) TEIJIN LTD.  
WPI; 1993-128866/16.  
Human protein C and activated protein C with short H chains -  
useful as anti-clotting agents and fibrinolysis promoters  
Disclosure; Fig 1; 8pp; Japanese.  
A human protein C or an activated protein C has a H chain contg. one  
of the residues 239-246 (= residues 450-457 in the sequence below)  
in the H chain of natural activated protein C as the C-terminal, or  
has a L chain contg. one of the residues 141-155 (= residues 141-155  
in the sequence below), pref. residues 149-155 (= residues 149-155  
in the sequence below), in the L chain of natural activated protein C  
as the C-terminal. The human protein C or the activated protein C  
as the C-terminal. The human protein C or a fibrinolysis promoter.

5Q	Sequence	461 AA;	
	Alignment Scores:	1.37e-159	
	Score:	2298.00	461
	Matches:	0	415
	Conservative:	0	
	Similarity:	100.00%	
	Best Local Similarity:	100.00%	
	Query Match:	98.33%	
	Misses:	0	
	Gaps:	0	

1-12451 x ABR34295 (1-461)

1	QY	GC	CAACTCTTCTGGAGGAGCTCCGTCACAGCAGCTGGAGCGGAGTGCATAGAGGAG	60
43	Db	Ala	AsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu	62
61	QY	AT	CTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATCTGGATGACACACTGGCCTTC	120
63	Db	Ile	CysAspPheGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe	82
121	QY	TG	TCTCAACACAGCTCGACGGTCAACAGTCCTGGCTCTGGCCCTTGGAGCACCCGTCGGCC	180
83	Db	Trp	SerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla	102
181	QY	AG	CCTGTGCTGGGGGACGGCACAGTCATCAGCGCATCGGCAGCTTCAGCTGCGACTGC	240
103	Db	Ser	LeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys	122
241	QY	CG	CAGCGCTGGAGGGCGCTCTCTGCCAGCGCAGGTGAGCTTCCTCAATTCGCTCCGCTG	300
123	Db	Arg	SerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu	142
301	QY	GAC	AACGGCGCTGCACGCAATCTACCTCTAGAGGAGTGGCTGGCGCGCTGAGCTGT	360
143	Db	Asp	AsnGlyGlyCysThrHisTyrCysLeuGluGluValGlyTrpArgArgCysSerCys	162

[illegible]

QY	361	GCCTCGGTACAAGTCGGGACGACCTCCTCGAGTGTACCCTCCGAGTGAAATGCCCT	420
Dd	163	AlaProGlyTyTLysLeuGlyAspAspLeuLeuCysHisProAlaValLysPhePro	182
QY	421	TGTGGAGCCCTGAAGCGGATGGAGAGAAGCGCACTCACCTGAAACGAGACACAGAA	480
Dd	183	CysGLyAArgProTrpPLysArgMetGLyLysLysArgSerHisLeuLysArgAspThrGlu	202
QY	481	GACCAAGAACCAAGTAGATCCCGCTCATTTGATGGAGATGACCGGGGGGAGAC	540
Dd	203	AspGlnGluAspGlnValAspProArgLeuLeuLeuAspGlyLysMetThrArgArgGlyAsp	222
QY	541	AGCCCTCGGAGGTCCTGCTGCTGACTCAAGAGAAGCTGGCTGCGGGGAGTGC	600
Dd	223	SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu	242
QY	601	ATCCACCCCTCTGGGTGCTGACAGCGGGCCACTGATGATGATGATGATGATGATGATG	660
Dd	243	IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGlySerLysLeuLeu	262
QY	661	GTCAGGCTTGAGAGATGACCTGCGGCTGGGAGAGTGGAGCTGGACCTGGACATC	720
Dd	263	ValArgLeuGlyGlyTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle	282
QY	721	AAGAGCTTCTGCTCCACCCCCTCACTACAGACAGACACACCGACATGATGATGATG	780
Dd	283	LysGluValPheValHisProAsnTrpSerLysSerThrThrAspAsnAspIleAlaLeu	302
QY	781	CTGACCTGCGCCCGCCGACCTCTGCGAGACCATAGTCCCATCTGCTCCCGGAC	840
Dd	303	LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp	322
QY	841	AGCGCTTTCAGAGCGGAGCTCAATCAGCGCGGCGGAGAGACCTCGTGACGGCTGG	900
Dd	323	SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp	342
QY	901	GGTACCACAGCGGAG	960
Dd	343	GlyTrpHisSerSerArgGlyLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe	362
QY	961	ATCAGATTCCGTGGTCCCGCAATAGTGCAGGAGGTGATGAGCAACATGTTGTCT	1020
Dd	363	IleLysIleProValValProHisAsnGlySerGluValMetSerAsnMetValSer	382
QY	1021	GAGAACATGCTGTGCGGCGATCTCCGGGACCGGACCGGAGATGCTGCGAGGCGACAGT	1080
Dd	383	GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer	402
QY	1081	GGGGGGCCATGTGTCGCTCTTCCAGGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTG	1140
Dd	403	GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp	422
QY	1141	GGTGGGCTGCTGGCTCTTCAACTACGCGCTTACACCAAGTACACCAAGTACACCAAG	1200
Dd	423	GlyGlyGlyCysGlyLeuLeuHisAsnTrpGlyValTrpThrLysValSerArgTrpLeu	442
QY	1201	GACTGATCATGGGCATCATGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1260
Dd	443	AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys	457
RESULT	15		
AAW02600			
ID	AAW02600	standard; Protein: 461 AA.	
XX	XX		
AC	AAW02600;		
DT	05-NOV-1996	(first entry)	
XX	Human protein C.		
KW	Activated protein C; serine protease; thrombosis; thrombolytic; fibrinolytic; antithrombotic; blood clotting; therapy.		

Tue Mar 18 16:19:30 2003

FT Disulfide-bond /label= N-glycosylation\_site  
 FT 373 /note= "forms disulphide bond with Cys387"  
 FT 387 /note= "forms disulphide bond with Cys373"  
 FT 398 /note= "forms disulphide bond with Cys426"  
 FT 426 /note= "forms disulphide bond with Cys 398"

US5516650-A.

14-MAY-1996.

27-JUN-1985; 85US-0749600.

28-FEB-1989; 89US-0317205.

27-JUN-1985; 85US-0749600.

29-OCT-1986; 86US-0924462.

08-DEC-1987; 87US-0130370.

10-SEP-1990; 90US-0582131.

04-DEC-1992; 92US-0987532.

08-APR-1994; 94US-0225253.

(ZYMO ) ZYMOGENETICS INC.

Berkner KL, Foster DC, Murray MJ;

WPI; 1996-251006/25.

N-PSDB; AAT32795;

N-PSDB; AAT32796.

New DNA encoding modified forms of opt. activated protein C - and related transformed cells for prodn. of recombinant protein C for use e.g. as an anti-thrombotic agent

Example 1; Fig 2A-C; 34pp; English.

Human protein C (AAW02600) is a zymogen of a serine protease that plays an important role in the regulation of blood coagulation and the generation of fibrinolytic activity in vivo. It is synthesised in the liver and processed to a 2-chain molecule, which is itself converted to activated protein C. Protein C and activated protein C are useful in the treatment of thrombotic disorders. They can be produced e.g. in mammalian host cells using a cDNA clone (AAT32795) derived from Hep G2 cells. Variant protein C, modified to improve cleavage between the heavy and light chains of the circulating intermediate, can also be produced.

Sequence 461 AA;

Alignment Scores:  
 Seq. No.: 1.37e-159 Length: 461  
 2298.00 Matches: 415  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 98.33%  
 Indels: 0  
 Gaps: 0  
 DB:

US-09-763-153-1 (1-1245) x AAW02600 (1-461)

QY 1 GCCAACTCTCTCGAGGAGCTCGTCCAGCAGCCTGGAGCGGAGTGCATAGAGGAG 60  
 Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
 QY 61 ATCTGTGACTTCGAGGAGCGCCAGGAAATTTTCCAAAATGTGGATGACACTGGCCCTTC 120  
 Db 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
 QY 121 TGGTCCAAAGCAGCTCAGCGGTGACAGTGTCTTGGCTTGGCTTGGAGCAGCCCGTGGCC 180  
 Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValProLeuGluHisProCysAla 102

QY 181 AGCTGTGTGCTGGCGGACGGCAGCTGCATCGACGCGCTTCAGTCCGAGTGCAGTGC 240  
 Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
 QY 241 CGCAGCGCTGGAGGCGCTTCTCCAGCGCGAGTGCAGTCTTCTCAATGTCTCGCTG 300  
 Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
 QY 301 GACAACGCGCTGCACGCAATTAATCTGCTAGAGAGTGGCTGGCGGCTGTAGTGTGT 360  
 Db 143 AspAsnGlyGlyCysThrHisThrCysLeuGluGluValGlyTrpArgArgCysSerCys 162  
 QY 361 GCGCTGTGCTACAAGCTGGGCGACGCTCTCTCAGTGTACCCCGCAGTGAAGTTCCT 420  
 Db 163 AlaProGlyTyrLysLeuGlyAspLeuGlnCysHisProAlaValLysPhePro 182  
 QY 421 TGTGGAGGCGCTGGAGCGGATGGAGGAGGAGGAGTGCATCTGAGTGCATCCCGCAGTGAAGTTCCT 480  
 Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
 QY 481 GACCAAGAGACCAAGTAGATCCGCGCTCATTTGATGGAGATGACCGCGCGGAGAC 540  
 Db 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgGlyAsp 222  
 QY 541 AGCCCTGCGCAGTGTCTCTGCTGACTCAAGAGAGAGTGGCTGCGGCGGAGTGTCT 600  
 Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 242  
 QY 601 ATCCACCTCTCTGGTGTGTGACAGCGGCGCTCATCTGATGATGATCCAGAGTCTCT 660  
 Db 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu 262  
 QY 661 GTCAGGCTTGGAGATGATGACCTGCGCGCTGGGAGAGTGGAGTGGAGTGGAGTGG 720  
 Db 263 ValArgLeuGlyGlyTrpAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAsp 282  
 QY 721 AAGGAGTCTTCTGCTCCACCCCAACTACAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
 Db 283 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 302  
 QY 781 GTGACCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840  
 Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
 QY 841 AGCGGCTTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900  
 Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
 QY 901 GGCTACACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
 Db 343 GlyTyrHisSerSerArgGlyLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
 QY 961 ATCAAGATTCCCGTGGTCCCGCACAAATGATGTCAGCGGAGTGCATGAGCAACATGGTGTCT 1020  
 Db 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
 QY 1021 GAGAACATCTGTGTCGGGCGATCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080  
 Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
 QY 1081 GGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140  
 Db 403 GlyClyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
 QY 1141 GGTGAGGCGTCTGGGCTCTTCAAACTACGGCGGTTTACACAAAGTACGGCGGTACCTC 1200  
 Db 423 GlyGlyGlyCysGlyLeuLeuHisAsnTyrGlyValThrLysValSerArgTyrLeu 442  
 QY 1201 CACTGTGATCCATGGCGCATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245  
 Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457  
 RESULT 16

AAV49561  
 ID AAY49561 standard; Protein; 461 AA.  
 XX AC AAY49561;  
 XX DT 13-JAN-2000 (first entry)  
 XX DE Human lecithin cholesterol acyltransferase protein sequence.  
 XX KW Human; coding sequence polymorphism; vascular pathology gene;  
 KW polymorphic site; phenotype correlation; forensic; paternity testing;  
 KW medicine; genetic analysis; vascular disease.  
 XX OS Homo sapiens.  
 XX PN W09950454-A2.  
 XX DT 07-OCT-1999.  
 XX DT 26-MAR-1999; 99WO-US06473.  
 XX PR 01-APR-1998; 98US-0054272.  
 XX PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
 XX PI Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;  
 XX DR WPI: 1999-620066/53.  
 XX DR N-PSDB; AA232180.  
 XX PT Determination of polymorphisms in genes, especially those identifying  
 PT predisposition to vascular disease -  
 XX PS Disclosure; Fig 24; 134pp; English.  
 XX CC AA232159 to AA232194 represent reference alleles for specifically  
 CC claimed nucleic acid sequences from the present invention which comprise  
 CC polymorphic sites as given in a table in the specification, selected  
 CC from 92 single nucleotide polymorphisms in which the nucleotide at the  
 CC polymorphic site is different from a nucleotide at the same site in a  
 CC reference allele. The nucleic acids, and primers and probes, are used to  
 CC identify polymorphisms, which may predispose an individual to disease,  
 CC especially a vascular disease. They can also be used in phenotype  
 CC correlations, forensics, paternity testing, medicine or genetic  
 CC analysis. AAY49550 to AAY49573 represent the proteins which correspond  
 CC to some of the reference alleles.

Sequence 461 AA;  
 Alignment Scores:  
 Pred. No.: 1.37e-159 Length: 461  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 20 Gaps: 0

US-09-763-153-1 (1-1245) x AAY49561 (1-461)  
 QY 1 GCCAATCTCTCTGAGGAGCTCGTCCACAGCAGCCTGGAGCGGAGTGCATAGAGGAG 60  
 Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerLeuGluArgGluGluGlu 62  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGATGACACACTGGCCTTC 120  
 Db 63 IleCysAspPheGluGluAlaLysGluLeuPheGlnAsnValAspThrLeuAlaPhe 82  
 QY 121 TGGTCCAGCAGCTGCAGCGTGACCTGCTTGGTCTTGGCCCTTGGAGCACCCTGGCGCC 180  
 Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
 QY 181 AGCTGTGCTCGGGGACGGCAGCTGCATCGACGCATCGGAGCTTCAGCTGCACCTGC 240

Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
 QY 241 CGCAGCGGCTGGGAGGGCGCTTCGCCAGCGCGAGGTGAGCTTCCTCAATGCTCGCTG 300  
 Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
 QY 301 GACAAGCGGCTGCACGCATTACTGCTAGAGGAGTGGGCTGGCGCGCTGAGCTGT 360  
 Db 143 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 162  
 QY 361 GGCCTTGGCTACAAGCTGGGGGAGCAGCTCTGAGTGTACCCCGCAGTGAATGTCCT 420  
 Db 163 AlaProGlyTyrLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 182  
 QY 421 TGTGGAGGCGCTGGAAGCGGATGGAAGAAGCGCAGTCACTGATGGGAGATGACGAG 480  
 Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
 QY 481 GACCAAGAAGACCAAGTAGATCCCGGCTCATTGATGGGAGATGACGAGGGGAGAC 540  
 Db 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
 QY 541 AGCCCTGGCAGGTGCTCTGCTGACTCAAGAAGAAGCTGGCTGGCGGCACTGCTC 600  
 Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 242  
 QY 601 ATCCACCCCTCCTGGGTGCTGACAGCGGCCCTGCTGATGATGATGATGATGATGAT 660  
 Db 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu 262  
 QY 661 GTCAGGCTTGAGAGTATGACCTGCGGCTGGGAGAGTGGGAGCTGGAGCTGGACATC 720  
 Db 263 ValArgLeuGlyGluTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
 QY 721 AAGAGGTCTTCTGTCACCCCACTACAGAACAGCACCACCAATGATGATGATGATGAT 780  
 Db 283 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 302  
 QY 781 CTGCACTTGGCCAGCGCCACCTCTCGCAGACCATAGTGGCCCATCTGCTCCCGGAC 840  
 Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
 QY 841 AGCGGCTTGCAGCGCGAGCTCAATCAGCCGCGCCAGGAGACCCCTGCTGACGGCTGG 900  
 Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
 QY 901 GCTACACAGCAGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 Db 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
 QY 961 ATCAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
 Db 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetSer 382  
 QY 1021 GAGAACATGCTGTGCGGCGATCCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080  
 Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
 QY 1081 GGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1140  
 Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
 QY 1141 GGTGAGGCTGTGGGCTCTTACAACTACGGCGGTTTACCAAAAGTCAGCGCTACCTC 1200  
 Db 423 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 442  
 QY 1201 GACTGATCCATGGGCATCATCAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245  
 Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

RESULT 17  
 AAE08626  
 ID AAE08626 standard; Protein; 461 AA.

Tue Mar 18 16:19:30 2003

XX AC AAE08626;  
 XX 01-NOV-2001 (first entry)  
 XX DE Human wild type protein C.  
 XX  
 XX Human: protein C derivative; anticoagulation activity; thrombosis;  
 KW serpin inactivation; acute coronary syndrome; myocardial infarction;  
 KW vascular occlusive disorder; hypercoagulable state; angina; sepsis;  
 KW disseminated intravascular coagulation; DIC; burn; transplantation;  
 KW sickle cell disease; viral haemorrhagic fever; protein C deficiency;  
 KW haemolytic uremic syndrome; acute arterial thrombotic occlusion;  
 KW thromboembolism; prothrombotic disorder; gene therapy; thalassaemia.  
 XX  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 XX Peptide 1-42  
 XX /label= Signal\_peptide  
 XX 43..461  
 XX /label= Mature\_human\_wild\_type\_protein\_C  
 XX  
 XX W0200159084-A1.  
 XX  
 XX 16-AUG-2001.  
 XX  
 XX 02-FEB-2001; 2001WO-US01221.  
 XX  
 XX 11-FEB-2000; 2000US-0181948.  
 XX  
 XX 14-MAR-2000; 2000US-0189199.  
 XX  
 XX (ELIL) LILLY & CO ELI.  
 XX  
 XX Gerlitz BE, Grinnell BW, Jones BE;  
 XX WPI; 2001-514662/56.  
 XX N-PSDB; AAD15224.  
 XX  
 XX Protein C derivative for treating acute coronary syndromes, vascular  
 XX occlusive disorders, thrombotic disorders and sepsis, comprises  
 XX substitutions at specified amino acid positions -  
 XX  
 XX Disclosure; Page 44-46; 59pp; English.  
 XX  
 XX The invention relates to human protein C derivatives and nucleic acid  
 XX molecules encoding such derivatives. These derivatives have increased  
 XX anticoagulation activity, resistance to serpin inactivation and  
 XX increased sensitivity to thrombin activation compared to wild type  
 XX protein C, and retains the biological activity of the wild type human  
 XX protein C. protein C derivatives are useful in the manufacture of a  
 XX medicament for the treatment of acute coronary syndromes e.g. myocardial  
 XX infarction and unstable angina; and disease states predisposing to  
 XX thrombosis; vascular occlusive disorders and hypercoagulable states e.g.  
 XX disseminated intravascular coagulation (DIC), burns, transplantations,  
 XX thalassaemia, sickle cell disease, viral haemorrhagic fever and  
 XX haemolytic uremic syndrome; sepsis in combination with bacterial  
 XX permeability increasing protein; thrombotic disorders in combination  
 XX with an anti-platelet agent; protein C deficiency; acute arterial  
 XX thrombotic occlusion, thromboembolism or stenosis in coronary, cerebral  
 XX or peripheral arteries or in vascular grafts in combination with a  
 XX thrombolytic agent. Nucleic acid molecules of the invention are useful  
 XX for treating humans with genetically predisposed prothrombotic disorders  
 XX by gene therapy. The present sequence is human wild type protein C.  
 XX  
 XX SQ Sequence 461 AA;  
 XX  
 XX Alignment Scores:  
 XX Pred. No.: 1.37e-159 Length: 461  
 XX Score: 2298.00 Matches: 415  
 XX Percent Similarity: 100.00% Conservative: 0  
 XX Best Local Similarity: 100.00% Mismatches: 0  
 XX Query Match: 98.33% Indels: 0

DB: 22 Gaps: 0  
 US-09-763-153-1 (1-1245) x AAE08626 (1-461)  
 QY 1 GCCAACTCTCTCTGGAGGAGCTCCGTACAGCAGCTGGAGCGGGAGTGCATAGAGGAG 60  
 DB 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluGluGluGlu 62  
 QY 61 ATCTGTGACTTCGAGGAGCCAGGAATTTCCAAAATGTGGATGACACACTGGCTTC 120  
 DB 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
 QY 121 TGCTCAACGACGTCGACGGTGCACAGTCTGTGGTCTTGGTCTGGAGCACCCTGGCC 180  
 DB 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
 QY 181 AGCTGTGTGGGGCAGCGCAGCTGCATCGACGCATCGCAGCTTCAGCTGGCATGC 240  
 DB 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
 QY 241 CGCAGCGCTGGAGGGCGCTTCTGCCAGCGGAGGTGAGCTTCTCAATTGCTCGCTG 300  
 DB 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
 QY 301 GACAAACGGGCTGCACGCACTTACTGCTAGAGAGGTGGCTGGCGGCTGTAGCTGT 360  
 DB 143 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 162  
 QY 361 GCGCTGGCTACAAAGCTGGGGGAGCAGCTCTCAGTGTACCCCGCAGTGAAGTTCCT 420  
 DB 163 AlaProGlyTyrLysLeuGlyAspLeuGlnCysHisProAlaValLysPhePro 182  
 QY 421 TGTGGAGGCGCTGGAGCGGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480  
 DB 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
 QY 481 GACCAAGAGACCAAGTAGATCCGCGCTCATGATGGGAGATGACAGGGGGGAGAC 540  
 DB 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
 QY 541 AGCCCTTGGCAGGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 DB 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242  
 QY 601 ATCCACCCCTCTCTGGGTGTGTGACAGCGCCCTGCTGCTGCTGCTGCTGCTGCT 660  
 DB 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLeuLeu 262  
 QY 661 GTGAGCTTGGAGAGTATGACCTGCGCGCTGGGAGAGTGGGAGCTGGAGCTGGAGATC 720  
 DB 263 ValArgLeuGlyGluTyrAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
 QY 721 AAGGAGGTCTCTGCTCCACCCCACTACAGCAGACACCCAGCAGCAATGACATCGCACTG 780  
 DB 283 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 302  
 QY 781 CTGACCTTGGCCAGCGCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
 DB 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
 QY 841 AGCGGCTTTCGAGAGCGAGCTCAATFAGCGCGCGGAGAGAGCCCTGCTGACGGCTGG 900  
 DB 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGlnGlnGlnGlnGlnGln 342  
 QY 901 GGTACACACAGCAGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 DB 343 GlyTyrHisSerSerArgGluLysGluLysLysLysLysLysLysLysLysLysLys 362  
 QY 961 ATCAAGATTCCCGTGGTCCCGCACAATAGTCAGGAGGAGTGCATGAGCAACATGTGTCT 1020  
 DB 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
 QY 1021 GAGAACATGCTGTGTGGGGCATCTCTCGGGGACCGGAGGATGCTCGGAGGCGGACGT 1080



Db 383 GluAsnMetLeuCySAlaGlyLeuLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
 QY 1081 GSGGGCCCATGGCTCCTTCCACGGCACCTGGTCTGGTGGCTGGTGGCTGG 1140  
 Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
 QY 1141 GGTGAGGCTGTGGGCTCCTTCAACTACACTGCGGCTTTACCAAGTCAGCGCTACCTC 1200  
 Db 423 GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValThrThrLysValSerArgTrpLeu 442  
 QY 1201 GACTGGATCCATGGGCACATCATGACAGACAGGAAGGCCCCACAGAG 1245  
 Db 443 AspTrpLeuHisGlyHisIleArgAspLysGluAlaProGlnLys 457

RESULT 18  
 AAB82674 standard; Protein; 461 AA.  
 AAB82674;  
 15-OCT-2001 (first entry)  
 Wild-type human protein C.  
 Protein C; human; coronary syndrome; thrombosis; angina;  
 myocardial infarction; vascular occlusive disorder;  
 hypercoagulation; sepsis; protein C deficiency; occlusion;  
 thromboembolism; stenosis; antibacterial; immunosuppressive;  
 thrombolytic; cardiac; antiangiinal; anticoagulant; therapy.  
 OS Homo sapiens.

Key Location/Qualifiers  
 Peptide 1..42  
 Protein /label= Signal\_peptide 43..461  
 Domain /label= Mature\_protein 43..87  
 Disulfide-bond /note= "Gla domain" 92..111  
 Disulfide-bond 101..106  
 Disulfide-bond 122..131  
 Disulfide-bond 140..151  
 Disulfide-bond 162..175  
 Disulfide-bond 183..319  
 Disulfide-bond 238..254  
 Disulfide-bond 373..387  
 Disulfide-bond 398..426  
 Cleavage-site 198..199  
 /note= "cleavage makes a 2-chain inactive precursor (155-amino acid light chain attached via a disulfide bond to a 262-amino acid heavy chain)"

Modified-site 48 /note= "gamma-carboxylated"  
 Modified-site 49 /note= "gamma-carboxylated"  
 Modified-site 56 /note= "gamma-carboxylated"  
 Modified-site 58 /note= "gamma-carboxylated"  
 Modified-site 61 /note= "gamma-carboxylated"  
 Modified-site 62 /note= "gamma-carboxylated"  
 Modified-site 67 /note= "gamma-carboxylated"  
 Modified-site 68 /note= "gamma-carboxylated"  
 Peptide 200..211 /note= "activation peptide; removal activates the 2-chain zymogen"

Cleavage-site 211..212 /note= "thrombin cleavage site"  
 Modified-site 139 /note= "N-glycosylated"  
 Modified-site 290 /note= "N-glycosylated"  
 Modified-site 355 /note= "N-glycosylated"  
 Modified-site 371 /note= "N-glycosylated"  
 Modified-site 44 /note= "N-glycosylated"  
 Modified-site 44 /note= "O-phosphorylated"  
 WO200157193-A2.  
 09-AUG-2001.  
 19-JAN-2001; 2001WO-US00020.  
 02-FEB-2000; 2000US-0179801.  
 14-MAR-2000; 2000US-0189197.  
 (ELIL) LILLY & CO ELI.  
 Gerlitz BE, Jones BE;  
 WPT; 2001-496919/54.  
 N-PSDB; AAB26362.  
 Novel human protein C derivative for treating, e.g., myocardial infarction, unstable angina, sepsis, thrombotic disorders, acute arterial thrombotic occlusion, and thromboembolism - Disclosure; Page 50-52; 63pp; English.

The present sequence is that of human protein C prepro-polypeptide. The invention relates to human protein C derivatives having at least 2 amino acid substitutions, and to recombinant DNA molecules encoding such derivatives. These derivatives have increased anticoagulant activity and resistance to inactivation by serpins compared with wild-type human protein C but retain the biological activity of the wild-type protein. The amino acid substitutions are selected from H10Q, S11G, S12K, Q32E, N33D, N33F, and amino acids at positions 194, 195, 228, 249, 254, 302, or 316 of the mature protein C. Polypeptide substituted with Ser, Ala, Thr, His, Lys, Leu, Arg, Asn, Asp, Glu, Gly or Gln (numbering relative to the protein C mature protein sequence). Preferred protein C derivatives are given in AAB82675-78. Also claimed are a vector comprising DNA encoding the novel human protein C derivatives, transformed host cells and a method of producing the human protein C derivatives. The protein C derivatives are useful for treating coronary syndromes and disease states predisposing to thrombotic disorders and infarction and unstable angina, vascular occlusive disorders and hypercoagulable states, sepsis (in combination with bactericidal permeability increasing protein or with tissue factor pathway inhibitor), thrombotic disorders (in combination with an anti-platelet agent or by local delivery through an intracoronary catheter), protein C deficiency, acute arterial thrombotic occlusion, thromboembolism, or stenosis in coronary, cerebral or peripheral arteries or in vascular grafts. Human patients with genetically predisposed thrombotic disorders may be treated by gene therapy (all claimed).

Sequence 461 AA;  
 Alignment Scores:  
 Pred. No.: 1 37e-159 Length: 461  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 22 Gaps: 0

us-09-763-153-1-n2p.rag

Tue Mar 18 16:19:30 2003

1081 GGGGGCCCATGGTCTCTCCAGCGACCTGGTCTCTGGTGGGCTGTGAGCTGG 1140  
 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
 1141 GGTGAGGCTGGGCTCTCTCACAACCTACGGGCTTTACACCAAGCTACGCGCTACCTC 1200  
 423 GlyGlyGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 442  
 1201 GACTGGATCCATGGGCATCATCAGACACAAGCAAGCAAGCAAGCAAGCAAG 1245  
 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

RESULT 19

AAB36895 standard; Protein; 461 AA.

XX AAB36895;

XX 26-FEB-2001 (first entry)

XX Human protein C derivative 2.

XX Protein C: human; vascular occlusive; burn; transplantation;

XX deep vein thrombosis; sickle cell; thalassemia;

XX thrombotic disorders; myocardial infarction; angina; stroke.

XX Homo sapiens.

XX WO2000066754-A1.

XX 09-NOV-2000.

XX 13-APR-2000; 2000WO-US08722.

XX 30-APR-1999; 99US-0131801.

XX (ELIL) LILLY & CO ELI.

XX Gerlitz BE, Jones BE;

XX WPI; 2001-007227/01.

XX N-PSDB; AAC83312.

protein C derivatives, useful for treating vascular occlusive disorder,  
 hypercoagulable state, thrombotic disorder and disease states  
 predisposing thrombosis, comprises specific amino acid substitutions -

Claim 1: Page 44-46; 57pp; English.

The present invention relates to a human protein C derivative. The  
 protein is useful for treating vascular occlusive disorders,  
 hypercoagulable states such as sepsis, disseminated intravascular  
 coagulation, purpura fulminans, major trauma, major surgery, burns,  
 adult respiratory distress syndrome, transplantation, deep vein  
 thrombosis, heparin-induced thrombocytopenia, sickle cell disease,  
 thalassemia, viral hemorrhagic fever, thrombotic thrombocytopenic  
 purpura, and hemolytic uremic syndrome, and also useful for treating  
 thrombotic disorders and acute coronary syndromes such as myocardial  
 infarction, unstable angina, and stroke. Protein C derivatives with  
 amino acid substitutions result in increased resistance to  
 inactivation by serpins when compared to wild-type activated human  
 protein C. They also have longer half-lives in human blood and hence  
 require either less frequent administration and/or smaller dosage  
 than wild type human protein C for treating disorders.

XX Sequence 461 AA;

Alignment Scores: 1.37e-159 Length: 461  
 Pred. No.: 2298.00 Matches: 415  
 Score: 100.00% Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00%

US-09-763-153-1 (1-1245) x AAB82674 (1-461)

1 GCCAATCTCTCTGAGGAGTCCGTCACACAGCCCTGGAGCGGAGTGCATAGAGGAG 60  
 43 AlaAsnSerPheLeuGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
 61 ATCTGTGACTTCAGAGGCGCAAGGAATTTCCAAATTTGGATGACACACTGGCCTTC 120  
 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
 121 TGGTCCAGACGCTCGACGCTGACAGTCTGGTCTTCCCTTGGAGCACCCGCTGGCC 180  
 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
 181 AGCTGTGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240  
 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
 241 CGCAGCGCTGGGAGGCGGCTTCTGCGGCGGAGGTGAGTCTCTCAATTGCTCGCTG 300  
 123 ArgSerGlyTrpGluGluArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
 301 GACAAGCGGCTGACGACGATTAATCTAGAGGAGTGGGCGGCGGCGGCGGCGGCGG 360  
 143 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 162  
 361 CGCCTGGCTACAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420  
 163 AlaProGlyTyrLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 182  
 421 TGTGGAGGCGGCTGGAAGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 183 CysGlyArgProIleLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
 481 GACCAAGAGACCAAGTAGATCGCGGCTCATGATGGAAGATGACAGCGGCGGAGAC 540  
 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
 541 AGCCCTGCGAGTGGTCTCTGCTGACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 223 SerProIleGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242  
 601 ATCCACCCCTCTGGGTGCTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660  
 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 262  
 661 GTACGCTGGAGAGTATGACCTGGCGGCGGCTGGAGAGTGGAGCTGGAGCTGGACATC 720  
 263 ValArgLeuGlyGluTyrAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
 721 AAGGAGTCTTCTCCACCCCACTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
 283 LysGluValPheValHisProAsnTyrSerLysSerThrAspAsnAspIleAlaLeu 302  
 781 CTGACCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840  
 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
 841 AGCGGCTTGGAGCGGAGCTCATCTAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900  
 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
 901 GGTACACAGCGGCGGAG 960  
 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
 961 ATCAAGATCCGCTGCTCCGACATATGAGTGCAGCGGAGTGCATGAGCAACATGTGTCT 1020  
 363 IleLysIleProValIleProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
 1021 GAGAACATGCTGTGTGGGCGGATCTCGGGAGCGGAGGAGTGCCTGCGAGGCGGACAGT 1080  
 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402

Query Match:	98.33%	Indels:	0
DB:	22	Gaps:	0
US-09-763-153-1 (1-1245) x AAB36895 (1-461)			
QY	1	GCACACCTCTCTCTGGAGGAGCTCCACACAGCCTGGAGCGGAGTGCATAGAGGAG	60
Db	43	AlaAsnSerPheLeuGluLeuArgHisSerLeuGluAlaGluCysIleGluGlu	62
QY	61	ATCTGTGACATCCAGAGGCCAAGGAAATTTCCAAATATGAGTACACACTGGCCTTC	120
Db	63	IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe	82
QY	121	TGTTCAAGACCTGCAGCGTACCAGTGTCTGCTGGCCCTGGAGCACCGCTGC	180
Db	83	TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla	102
QY	181	AGCTGTGCTGGCGCACGACCTGATCGAGCGATCGGAGTGGCTGGCGCTGAGTGT	240
Db	103	SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys	122
QY	241	CGACGCTGGAGGCGCTCTCTGCGAGCGAGTGGCTGGCGCTGAGTGT	300
Db	123	ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu	142
QY	301	GACACGCGGCTGCAGCATTACTCCCTAGAGAGTGGCTGGCGCTGAGTGT	360
Db	143	AspAsnGlyGlyCysThrHisThrCysLeuGluGluValGlyTrpArgA9CysSerCys	162
QY	361	CGCGCTGGCTACAGTGGGAGCAGCTCTCTGAGTGTACCCCGCAGTGAAGTTCCT	420
Db	163	AlaProGlyTrpLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro	182
QY	421	TGTTGGAGGCGCTGGAAGCGATGAGAGAGCGCAGTCACTGAAACGAGACAGAA	480
Db	183	CysGlyArgProTrpLysArgMetGluLysArgSerHisLeuLysArgAspThrGlu	202
QY	481	GACCAAGAACCAAGTAGTCCGCGCTCATTTGATGGGAAGATGACACGCGGGAGAC	540
Db	203	AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgGlyAsp	222
QY	541	AGCCCTGGAGTGGTCTGCTGGACTCAAGAACAGCTGGCTGGGCGCAGTGTCT	600
Db	223	SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu	242
QY	601	ATCCACCCCTCTCTGGTGTGACAGCGCGCCCTCATGATGATGATGATGATGATGAT	262
Db	243	IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu	282
QY	661	GTCAGGCTGGAGTATGACCTGGCGGCTGGAGAGTGGAGTGGAGTGGAGTGGAGT	322
Db	263	ValArgLeuGlyGlyTrpAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle	342
QY	721	AAGGAGTCTCTGTCACCCCACTACAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAG	382
Db	283	LysGluValPheValHisProAsnTrpSerLysSerThrAspAsnAspIleAlaLeu	402
QY	781	CTGCACCTGGCCCGCCCGCCCTCTCTGAGACCATGATGATGATGATGATGATGATG	442
Db	303	LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp	462
QY	841	AGCGGCTTGCAGAGCGAGCTCAATCAGCGCGCGCGCGCGCGCGCGCGCGCGCG	502
Db	323	SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp	522
QY	901	GGCTACACAGCAGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	562
Db	343	GlyTrpHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe	582
QY	961	ATCAGATTCCTGTCTCCCGCAGATGATGATGATGATGATGATGATGATGATGATG	622
Db	363	IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer	642
QY	1021	GAGAACATGCTGTGCTGGCGGATCTCTGGGAGCAGGAGGAGGAGGAGGAGGAGGAG	1080
Db	383	GluAsnMetLeuGlyAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer	1100
QY	1081	GGGGGGCCCATGCTCCCTCCCTCCACGCGACCTGGTTCCTGGTGGGCGCTGGTGG	1140
Db	403	GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp	1160
QY	1141	GSTGAGGCTGGGCTCTTCACACCTACACCTACACCTACACCTACACCTACACCTAC	1200
Db	423	GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValTrpThrLysValSerArgTrpLeu	1240
QY	1201	GACTGATCCATGGGCGACATCAGACAGACAGGAGGAGGAGGAGGAGGAGGAGGAG	1245
Db	443	AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys	1250
RESULT	20		
AAU99001	ID	AAU99001 standard; Protein; 461 AA.	
XX	AC	AAU99001;	
XX	DT	23-AUG-2002 (first entry)	
XX	DE	Human Protein C precursor protein.	
XX	KW	Human; Protein C; N-glycosylation; APC; activated protein C; precursor;	
XX	KW	serum half-life; chromosome 2q13-q14; stroke; myocardial infarction;	
XX	KW	sepsis; septic shock; disseminated intravascular coagulation; DIC;	
XX	KW	bone marrow transplantation; major surgery; trauma; ARDS; coagulant;	
XX	OS	Homo sapiens.	
XX	Key	Location/Qualifiers	
FT	Peptide	1..42	
FT	Protein	/label= Signal_peptide	
FT	Protein	43..461	
FT	Protein	/label= Mature_protein_C	
FT	Peptide	43..197	
FT	Peptide	/label= Light_chain	
FT	Peptide	198..199	
FT	Peptide	/label= Lys_Arg_dipeptide	
FT	Peptide	200..301	
FT	Protein	/label= Activation_peptide	
FT	Protein	200..461	
FT	Protein	/label= Heavy_chain	
XX	WO200232461-A2.		
XX	25-APR-2002.		
XX	15-OCT-2001; 2001WO-DK00679.		
XX	18-OCT-2000; 2000DK-0001560.		
XX	18-OCT-2000; 2000US-242268P.		
XX	21-JUN-2001; 2001DK-0000970.		
XX	21-JUN-2001; 2001US-300154P.		
XX	(MAXY-) MAXYGEN APS.		
XX	(MAXY-) MAXYGEN HOLDINGS LTD.		
XX	Andersen KV, Pedersen AH, Freskgaard PO;		
XX	WPI; 2002-489875/52.		
XX	N-PSDB; ABR86038.		
XX	Novel conjugate useful for treating or preventing septic shock, stroke		
XX	and myocardial infarction, comprises non-polypeptide group covalently		
XX	attached to protein C polypeptide comprising an attachment group -		
XX	Example 4; Page 76-77; 92pp; English.		

[illegible]

QY  
Db

121 TGGTCCAGGCGCTTT  
83 TrpSerLysHisValAspGlyAs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run On: March 17, 2003, 17:26:59 ; Search time 51 Seconds

(without alignments)

4693.622 Million cell updates/sec

Title: US-09-763-153-1

Perfect score: 2337

Sequence: 1 gccaaactctctctggagga.....acaaggaagccccccagaag 1245

ring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 60 summaries

Command line parameters:

-MODEL=frame+n2p\_model -DEV=xlp  
-O=/cgn2.1/USPTO\_spoel/US09763153/runat\_11032003\_084248\_2005/app\_query.fasta\_1.1415  
-DB=PIR\_73 -QFMT=fastcan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=60  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=20 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09763153@cgn.1.1.38 @runat\_11032003\_084248\_2005 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2298	98.3	461	1 KXHU	protein C (activat
2	1660	71.0	456	1 KXBO	protein C (activat
3	1650.5	70.6	461	1 S18994	protein C (activat
4	1637.5	70.1	461	1 JX0210	protein C (activat
5	818.5	35.0	482	1 EXRT	coagulation factor
6	809	34.6	482	1 EXBO	coagulation factor
7	801.5	34.3	475	1 EXCH	coagulation factor
8	801.5	34.3	488	1 EXHU	coagulation factor
9	783	33.5	466	1 KFHU7	coagulation factor
10	779.5	33.4	407	1 KFB07	coagulation factor
11	769.5	32.9	443	2 I46932	coagulation factor
12	763	32.6	432	1 A30351	coagulation factor
13	736	31.5	461	1 KFHU	coagulation factor
14	726	31.1	459	2 JQ0419	coagulation factor

15	714.5	30.6	416	1 KFBO	coagulation factor
16	562.5	24.1	622	1 TBHU	thrombin (EC 3.4.2
17	538.5	23.0	625	1 TBBO	thrombin (EC 3.4.2
18	533	22.8	618	2 A35827	thrombin (EC 3.4.2
19	525.5	22.5	617	2 S10511	thrombin (EC 3.4.2
20	470	20.1	655	1 A46688	hepatocyte growth
21	468	20.0	422	1 KXHU2	plasma protein 2 p
22	445.5	19.1	396	1 KXBOZ	plasma protein 2 -
23	427	18.3	271	2 I46580	factor IX - piq (f
24	426	18.2	638	1 KQMSPL	plasma kallikrein
25	424.5	18.2	638	1 KQHUP	plasma kallikrein
26	420.5	18.0	699	1 I54763	Ra-reactive factor
27	420	18.0	27	2 I46712	factor IX - rabbit
28	419.5	18.0	812	1 PLMS	plasma (EC 3.4.21
29	419	17.9	625	1 KFHU1	coagulation factor
30	418.5	17.9	855	2 JC7731	membrane-bound arg
31	418.5	17.9	855	2 JC7775	membrane type-sei
32	417	17.8	285	2 I48144	coagulation factor
33	414	17.7	638	1 KQRTPL	plasma kallikrein
34	411	17.6	786	1 A47547	serine proteinase
35	410.5	17.6	275	2 C35863	trypsin (EC 3.4.2
36	410	17.5	560	1 JC4795	plasma hyaluronan-
37	410	17.5	810	1 PLHU	plasma (EC 3.4.21
38	409.5	17.5	282	2 I84621	coagulation factor
39	406	17.4	417	1 S00845	hepsin (EC 3.4.21
40	404.5	17.3	275	2 A35863	trypsin (EC 3.4.2
41	404	17.3	1019	2 A38738	coagulation factor
42	403.5	17.3	275	2 B35863	trypsin (EC 3.4.2
43	403	17.2	583	2 A29154	trypsin (EC 3.4.2
44	402	17.2	246	2 B25528	complement factor
45	400.5	17.1	246	2 JQ1471	trypsin (EC 3.4.21
46	400	17.1	239	2 G42696	thrombin (EC 3.4.2
47	400	17.1	810	2 B30848	plasma (EC 3.4.21
48	399.5	17.1	231	1 TRFCTR	trypsin (EC 3.4.21
49	399	17.1	274	2 JC4171	trypsin (EC 3.4.2
50	399	17.1	1019	1 A56318	enteropeptidase (E
51	398	17.0	235	2 D42696	thrombin (EC 3.4.2
52	398	17.0	1047	2 A56117	masquerade precurs
53	397.5	17.0	236	2 C42696	thrombin (EC 3.4.2
54	396.5	17.0	270	2 S56160	mast cell tryptase
55	396.5	17.0	275	2 A32410	trypsin (EC 3.4.2
56	396.5	17.0	276	2 A38654	trypsin (EC 3.4.2
57	396	16.9	235	2 H42696	thrombin (EC 3.4.2
58	395	16.9	246	1 TRFT2	trypsin (EC 3.4.21
59	392.5	16.8	251	2 PC1235	29K serine protein
60	392	16.8	264	2 I38136	chymotrypsin-like

# ALIGNMENTS

## RESULT 1

KXHU

protein C (activated) (EC 3.4.21.69) precursor - human

N/Alternate names: autoprothrombin IIA; plasma protein C

C/Species: Homo sapiens (man)

C/Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 16-Jul-1999

C/Accession: A22331; A25426; A21781; A23789; A00927

R/Foster, D.C.; Yoshitake, S.; Davie, E.W.

A/Title: The nucleotide sequence of the gene for human protein C.

A/Reference number: A22331; MUID:85270390; PMID:2991887

A/Accession: A22331

A/Molecule type: DNA

A/Residues: 1-461 <FOS1>

A/Cross-references: GB:M11228; NID:g190333; PIDN:AAA60166.1; PID:g190334

R/Plutsky, J.; Hoskins, J.A.; Long, G.L.; Crabtree, G.R.

Proc. Natl. Acad. Sci. U.S.A. 83, 546-550, 1986

A/Title: Evolution and organization of the human protein C gene.

A/Reference number: A25426; MUID:86120978; PMID:3511471

A/Accession: A25426

A/Molecule type: DNA

A/Residues: 1-445, 'L', 446-461 <PLU>

A/Cross-references: GB:M12712; NID:g190330; PIDN:AAA60165.1; PID:g190332

R; Foster, D.; Davie, E.W.  
 Proc. Natl. Acad. Sci. U.S.A. 81, 4766-4770, 1984  
 A: Title: Characterization of a cDNA coding for human protein C.  
 A: Reference number: A21781; MUID: 84272714; PMID: 6589623  
 A: Accession: A21781  
 A: Molecule type: mRNA  
 A: Residues: 1-461 <FOS2>  
 A: Cross-references: GB:K02059; NID: g190322; PIDN: AAA60164.1; PID: g190323  
 R; Beckmann, R.J.; Schmidt, R.J.; Satterre, R.F.; Plutsky, J.; Crabtree, G.R.; Long, G.L.  
 Nucleic Acids Res. 13, 5233-5247, 1985  
 A: Title: The structure and evolution of a 461 amino acid human protein C precursor and its  
 A: Reference number: A23789; MUID: 85269639; PMID: 2991859  
 A: Accession: A23789  
 A: Molecule type: mRNA  
 A: Residues: 1-461 <BEC>  
 A: Cross-references: GB: X02750; NID: g35689; PIDN: CAA26528.1; PID: g763120  
 R; Mielich, J.P.; Broze Jr., G.J.  
 J. Biol. Chem. 265, 11397-11404, 1990  
 A: Title: Beta protein C is not glycosylated at asparagine 329. The rate of translation m  
 A: Reference number: A44605; MUID: 90293094; PMID: 1694179  
 A: Contents: annotation: carbohydrate binding sites: activation peptide  
 A: Comment: the alpha form of protein C is glycosylated at Asn-329, and the beta form is not  
 A: Title: O-linked fucose is present in the first epidermal growth factor domain of facto  
 A: Reference number: A44606; MUID: 92184750; PMID: 1544894  
 A: Comment: annotation: beta-hydroxyaspartic acid  
 A: Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that in  
 A: Comment: Protein C is strongly enhanced by complexing with protein S. Protein C also f  
 A: Comment: Protein C is synthesized in the liver as a single chain precursor, which is c  
 A: Comment: Protein C is a dodecapeptide from the amino end of the heavy chain; this reaction  
 C: Genetics:  
 A: Gene: GDB: PROC  
 A: Cross-references: GDB: 120317; OMIM: 176860  
 A: Map position: 2q13-2q21  
 A: Introns: 24/1; 79/3; 88/1; 134/1; 179/1; 226/3; 266/1  
 C: Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology  
 F: 1-32/Domain: anticoagulant; beta-hydroxyaspartic acid; blood coagulation; calcium binding  
 F: 1-32/Domain: signal sequence #status predicted <SIG>  
 F: 27-86/Domain: Gla domain homology <GLA>  
 F: 33-42/Domain: propeptide #status predicted <PRO>  
 F: 43-197/Product: protein C light chain #status predicted <LCH>  
 F: 92-131/Domain: EGF homology <EG1>  
 F: 140-175/Domain: EGF homology <EG2>  
 F: 200-461/Product: protein C heavy chain #status predicted <HCH>  
 F: 200-211/Domain: activation peptide #status experimental <APT>  
 F: 212-445/Domain: trypsin homology <TRY>  
 F: 48, 49, 56, 58, 61, 62, 67, 68, 71/Modified site: gamma-carboxyglutamic acid (Glu) #status exp  
 F: 59-64, 92-105, 101-120, 122-131, 140-151, 147-160, 162-175, 183-319, 238-254, 373-387, 398-426/1  
 F: 106-111/Disulfide bonds: #status predicted  
 F: 110/Binding site: carboxylate (Thr) (covalent) #status absent  
 F: 113/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental  
 F: 330, 335/Binding site: carboxylate (Asn) (covalent) #status experimental  
 F: 1-212/Cleavage site: Arg-Leu (thrombin) #status experimental  
 F: 239, 402/Active site: His, Asp, Ser #status predicted  
 F: 371/Binding site: carboxylate (Asn) (covalent) (partial) #status atypical

Alignment Scores:  
 Pred. No.: 2,41e-135 Length: 461  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 1 Gaps: 0

US-09-763-153-1 (1-1245) x KXHU (1-461)

QY 1 GCCAACTCCTCGAGAGCTCGCTACAGACAGCCCTGGAGCGGAGTGCATAGAGAG 60  
 Db 43 ALaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
 QY 61 APTCTGTGCTTCGAGAGCGCCAAAGAAATTTCCAAATGTGGATGACACACTGGCCTTC 120

Db 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
 QY 121 TGTGTCACAGACAGCTGACGGTGACAGTGTGTGGCTTGTCCCTTGGAGACACCCGTCGCC 180  
 Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
 QY 181 AGCCTGTGCTGCGGGCAGCGGACGTCATCGAGCGCATCGGAGCTTCAGTGGCAGTGC 240  
 Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
 QY 241 CCGACGGCTGGAGGGCGGCTTCTCCAGCGCGAGGTGAGCTTCTCAATGTCTCGCTG 300  
 Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
 QY 301 GACAAACGGCGCTGCACGACATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
 Db 143 AspAsnGlyGlyCysThrHisThrCysLeuGluGluValGlyTrpArgArgCysSerCys 162  
 QY 361 GCGCTGCTTACAAGCTGGGGGACGACCTCTGTCAGTGTGACCCCGCAGTGAAGTTCCT 420  
 Db 163 AlaProGlyTyrLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro 182  
 QY 421 TGTGGGAGCCCTGGAAGCGGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 Db 183 CysGlyArgProtrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
 QY 481 GACCAAGAACACAAAGTAGATCCGCGCTTATGTATGGAGAGATGACAGCGGGGAGAC 540  
 Db 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
 QY 541 AGCCCTGCGAGGTGTCTCTGTGACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242  
 QY 601 ATCCACCCCTCTGCTGGTGTGTGACAGCGGCCACCTGATGGATGAGTCCAAAGAGCTCT 660  
 Db 243 IleHisProSerTrpValLeuThrAlaHisCysMetAspGluSerLysLysLeuLeu 262  
 QY 661 GTCAGCTTGGAGATGATACCTCGGGCGCTGGGAGAGAGAGAGAGAGAGAGAGAGAG 720  
 Db 263 ValArgLeuGlyGluTyrAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
 QY 721 AAGGAGGTCTGTGTCACCCCACTACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
 Db 283 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 302  
 QY 781 CTCACCTGCGCCAGCCGCCACCTCTCGCAGACATAGTCCCATCTGCTCCCGGAC 840  
 Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
 QY 841 AGCGGCTTTCAGAGCGCGAGCTCATATGAGCGCGCGAGAGAGAGAGAGAGAGAGAG 900  
 Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
 QY 901 GGTACACAGACAGCGAG 960  
 Db 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
 QY 961 ATCAAGATTCCTGTCGCCGACAAATGATGAGCGAGGATGATGAGCAACATGTGTCT 1020  
 Db 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
 QY 1021 GAGAACATGCTGTGCGGGCATCTCGGGGAGCGCGAGAGAGAGAGAGAGAGAGAGAG 1080  
 Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
 QY 1081 GGGGGGCGCATGCTGCTCTTCCAGCGCACCTGTTCTCTGTTGGGCGCTGTGTGAGCTGG 1140  
 Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
 QY 1141 GGTGAGGCTGTGGGCTCTCTTCAACTACGCGGCTTTACACCAAGTACACCGGCTACCTC 1200  
 Db 423 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 442

QY 1201 GACTGGATCATGGGCACATCAGACAAAGGAAGCCGCCAGAAAG 1245  
 |||||  
 Db 443 ASPTripIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

## RESULT 2

KXBO

protein C (activated) (EC 3.4.21.69) precursor - bovine (fragment)  
 N;Alternate names: autoprothrombin IIA; plasma protein C  
 C;Species: Bos primigenius taurus (cattle)  
 C;Date: 30-Nov-1980 #sequence\_revision 17-Mar-1987 #text\_change 16-Jul-1999  
 C;Accession: A26250; A18385; A18386; A00928  
 R;Long, G.L.; Balagaje, R.M.; Macgillivray, R.T.A.  
 Proc. Natl. Acad. Sci. U.S.A. 81, 5653-5656, 1984  
 Title: Cloning and sequence of liver cDNA coding for bovine protein C.  
 Reference number: A26250; MUID:85014826; PMID:5091100

Accession: A26250

Molecule type: mRNA

A;Residues: 1-456 <LON>

R;Fennlund, P.; Stenflo, J.

J. Biol. Chem. 257, 12170-12179, 1982

A;Title: Amino acid sequence of the light chain of bovine protein C.

A;Reference number: A18385; MUID:83007325; PMID:6896876

A;Accession: A18385

A;Molecule type: protein

A;Residues: 40-194 <FER>

A;Note: 82-Lys was also found

R;Brakenberg, T.; Fennlund, P.; Roepstorff, P.; Stenflo, J.

Proc. Natl. Acad. Sci. U.S.A. 80, 1802-1806, 1983

A;Title: beta-Hydroxyaspartic acid in vitamin K-dependent protein C.

A;Reference number: A19316; MUID:83169769; PMID:6572939

A;Contents: annotation; revision to residue 110

R;Stenflo, J.; Fennlund, P.

J. Biol. Chem. 257, 12180-12190, 1982

A;Title: Amino acid sequence of the heavy chain of bovine protein C.

A;Reference number: A18386; MUID:83007326; PMID:6896877

A;Accession: A18386

A;Molecule type: protein

A;Residues: 197-454, 'pv' <STEX>

R;Esmon, N.L.; DeBault, L.E.; Esmon, C.T.

J. Biol. Chem. 258, 5548-5553, 1983

A;Title: Proteolytic formation and properties of gamma-carboxyglutamic acid-domainless F

A;Reference number: A37541; MUID:83213513; PMID:6304092

A;Contents: annotation; activation; calcium binding

R;Johnson, A.E.; Esmon, N.L.; Laue, T.M.; Esmon, C.T.

J. Biol. Chem. 258, 5554-5560, 1983

A;Title: Structural changes required for activation of protein C are induced by Ca2+ bin

A;Reference number: A37542; MUID:83213514; PMID:6406503

A;Contents: annotation; activation; calcium binding

C;Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that re

s.

C;Comment: Protein C is synthesized in the liver as a single chain precursor, which is c

bin, which cleaves a tetradecapeptide from the amino end of the heavy chain; this reacti

C;Comment: Calcium binds to the gamma-carboxyglutamic acid (Gla) residues and, with stro

cognition of the thrombin-thrombomodulin complex.

C;Comment: The gamma-carboxyglutamic acid residues arise by a posttranslational, vitamin

C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology

C;Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coagulation; calcium binding

F;1-29/Domain: signal sequence (fragment) #status predicted <SIG>

F;24-83/Domain: Gla domain homology <GLA>

F;80-194/Product: propeptide #status predicted <PRO>

F;98-128/Domain: EGF homology <EGF>

F;137-172/Domain: EGF homology <EG2>

F;197-456/Product: protein C heavy chain #status experimental <LCH>

F;197-210/Domain: activation peptide #status experimental <ACT>

F;211-440/Domain: trypsin homology <TRY>

F;45-46,53,55,58,59,62,64,65,68,74/Modified site: gamma-carboxyglutamic acid (Glu) #stat

F;110/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental

F;119-128,137-148,144-157,159-172,180-318,323-353,358-382,393-421/Disulfide bonds: #stat

F;136,289,350/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;252,298,397/Active site: His, Ser #status predicted

F;366/Binding site: carbohydrate (Asn) (covalent) #status predicted

## Alignment Scores:

Pred. No.: 1.08e-95 Length: 456  
 Score: 1660.00 Matches: 298  
 Percent Similarity: 80.96% Conservative: 38  
 Best Local Similarity: 71.81% Mismatches: 73  
 Query Match: 71.03% Indels: 6  
 DB: 1 Gaps: 2

US-09-763-153-1 (1-1245) x KXBO (1-456)

QY 1 GCCAACTCTCTCTGGAGAGCTCCGTACACAGCAGCCTGGAGCGGAGTGCATAGAGGAG 60  
 |||||  
 Db 40 AlaAsnSerPheLeuGluGluLeuArgProGlyAsnValGluArgGluCysSerGluGlu 59  
 QY 61 ATCTGTACTCTCAGGAGGCCAAGAAATTTCCAAATGTGGATGACACACTGGCCTTC 120  
 :::::|||||  
 Db 60 ValCysGluPheGluGluAlaArgGluPheGlnAsnThrGluAspThrMetAlaPhe 79  
 QY 121 TGGTCCAAAGCAGTCGACGGTGACAGTCTTGGTCTTGCCCTTGGAGACCCCTGGGCC 180  
 |||||  
 Db 80 TrpSerPheTyrSerAspGlyAspGlnCysGluAspArgProSerGlySerProCysAsp 99  
 QY 181 AGCCTGTCTCGGGCAGCGCAGCTGCATCGACGCGCATCGCAGCTTCAGCTCGCACTGC 240  
 |||||  
 Db 100 LeuProCysCysGlyArgGlyLysCysIleAspGlyLeuGlyPheArgCysAspCys 119  
 QY 241 CGCAGCGCTGGAGGGCGCTTCTGCCAGCGCGAGGTGAGCTTCCTCAATTTGCTCGCTG 300  
 |||||  
 Db 120 AlaGluGlyTyrGluGlyArgPheCysLeuHisGluValArgPheSerAsnCysSerAla 139  
 QY 301 GACACGGCGCTGCACCGCATCTACTGCTAGCAGAGGTGGCGCTGCTAGCTGT 360  
 :::::|||||  
 Db 140 GluAsnGlyGlyCysAlaHisTyrCysMetGluGluGlyArgArgHisCysSerCys 159  
 QY 361 CGCCTGTGCTACAAGCTGGGGCAGCAGCTCTCGCAGTGTCAACCCGCGAGTGAAGTCCCT 420  
 |||||  
 Db 160 AlaProGlyTyrArgLeuGluAspHisGlnLeuCysValSerLysValThrPhePro 179  
 QY 421 TGTGGAGCGCTCGAAGCGGATGGAGAAAGCGCAGTCACTCAAGAACGACAGACAGAA 480  
 |||||  
 Db 180 CysGlyArgLeuGlyLysArgMetGluLysLysArgLysThrLeuLysArgAspThrAsn 199  
 QY 481 -----GACCAAGAGACCAAGTAGATCCGCGGCTCATTTGATGGAGAGATCACCAGCGG 534  
 |||||  
 Db 200 GlnValAspGlnLysAspGlnLeuAspProArgIleValAspGlyGlnGluAlaGlyTyr 219  
 QY 535 GGAGACAGCCCTCGCAGGTGCTCTGCTGGACTCAAGAAAGAGAGCTGGCGCTCGGGGCA 594  
 |||||  
 Db 220 GlyGluSerProTyrGlnAlaValLeuLeuAspSerLysLysLeuValCysGlyAla 239  
 QY 595 GTGCTATCCACCCCTCTCTGGTGTGCTGACACGGGCCCGCAGTGCATGGATGAGTCAAGAG 654  
 |||||  
 Db 240 ValLeuIleHisValSerTyrValLeuThrValAlaHisCysLeuAspSerArgLysLys 259  
 QY 655 CTCCTTCTCAGCTTGGAGATGATGACCTGGCGGCTGGGAGAGAGTGGAGCTGGACCTG 714  
 |||||  
 Db 260 LeuIleValArgLeuGlyLysTyrAspMetArgTyrGluSerTyrGluValAspLeu 279  
 QY 715 GACATCAAGGAGGTCTTCTGCCACCCCACTACAGAGAGAGCAGCAGCAATGACATC 774  
 |||||  
 Db 280 AspIleLysGluValIleIleHisProAsnTyrThrLysSerThrSerAspAsnIle 299  
 QY 775 GCATGTGTGACCTGGGCCCGCCAGCCCTCTCGCAGACCATAGTGGCCCATCTGCTC 834  
 |||||  
 Db 300 AlaLeuLeuArgLeuAlaLysProAlaThrLeuSerGlnThrIleValProIleCysLeu 319  
 QY 835 CCGGACAGCGGCTTGCAGAGCGGAGCTCAATCAGCGCCGCCAGAGACCTCTGTCAGC 894  
 |||||  
 Db 320 ProAspSerGlyLeuSerGluArgLysLeuThrGlnValGlyGlnGluThrValValThr 339  
 QY 895 GGCTGGGGCTACACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 954  
 |||||









Db 385 ArgAsnThrCysArgLeuSerThrSerPheSerIleThrGlnAsnMetPheCysAlaGly 404  
 QY 1042 ATCTCGGGACCGGAGGATCCCTCGAGGGGACAGTGGGGGCCCATGGTCGCCCTCC 1101  
 Db 405 TyrAspAlaLysGlnGluAspAlaCysGlnGlyAspSerGlyGlyProHisValThrArg 424  
 QY 1102 TTCCAGCGGACCTGGTCTCTGGTGGCCCTGGTGGAGCTGGGTGGGGCTGCTCTT 1161  
 Db 425 PheLysAspThrTyrPheValThrGlyIleValSerTrpGlyGluGlyCysAlaArgLys 444  
 QY 1162 CACACTACGGGTTTACACCAAGTCACCGCTACCTCGACTGGATCCATGGGCACATC 1221  
 Db 445 GlyLysTyrGlyIleTyrThrLysValThrAlaPheLeuLysTrpIleAspArgSerMet 464  
 1222 AGAGACAGGAAGCCGCC 1239  
 465 LysAlaArgValGlyPro 470

## RESULT 6

## EXBO

coagulation factor Xa (EC 3.4.21.6) precursor - bovine  
 N:Alternate names: Stuart factor  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 24-Apr-1984 #sequence\_revision 17-Mar-1987 #text\_change 16-Jul-1999  
 C:Accession: A22867; A14997; A12030; A34412; S39414; A00925  
 R:Fung, M.R.; Campbell, R.M.; MacGillivray, T.A.  
 Nucleic Acids Res. 12, 4481-4492, 1984  
 A:Title: Blood coagulation factor X mRNA encodes a single polypeptide chain containing a  
 A:Reference number: A22867; MUID:84247315; PMID:6330671  
 A:Accession: A22867  
 A:Molecule type: mRNA  
 A:Residues: 1-487 <FUN>  
 A:Cross-references: GB:X00673; NID:g192; PIDN:CAA25286.1; PID:g193  
 R:Enfield, D.L.; Ericsson, L.H.; Fujikawa, K.; Walsh, K.A.; Neurath, H.; Titani, K.  
 Biochemistry 19, 659-667, 1980  
 A:Title: Amino acid sequence of the light chain of bovine factor X-1 (Stuart factor).  
 A:Reference number: A14997; MUID:80130563; PMID:6766735  
 A:Accession: A14997  
 A:Molecule type: protein  
 A:Residues: 41-102, N', 104-180 <ENF>  
 R:McMullen, B.A.; Fujikawa, K.; Kisiel, W.  
 Biochem. Biophys. Res. Commun. 115, 8-14, 1983  
 A:Title: The occurrence of beta-hydroxyaspartic acid in the vitamin K-dependent blood co  
 A:Reference number: A20274; MUID:83308813; PMID:6688526  
 A:Contents: annotation; revision to residue 103  
 R:Titani, K.; Fujikawa, K.; Enfield, D.L.; Ericsson, L.H.; Walsh, K.A.; Neurath, H.  
 Proc. Natl. Acad. Sci. U.S.A. 72, 3082-3086, 1975  
 A:Title: Bovine factor X-1 (Stuart factor): amino-acid sequence of heavy chain.  
 A:Reference number: A12030; MUID:76053069; PMID:1059093  
 A:Accession: A12030  
 A:Molecule type: protein  
 A:Residues: 183-252, 294-295, 'GBE', 299-334, 336-348, 'AE', 351-354, 356-441, 'GRFG', 446-492 <T  
 A:Note: carbohydrate binding sites and disulfide bonds were determined  
 R:Persson, E.; Selander, M.; Linse, S.; Drakenberg, T.; Oehlén, A.K.; Stenflo, J.  
 J. Biol. Chem. 264, 16897-16904, 1989  
 A:Title: Calcium binding to the isolated beta-hydroxyaspartic acid-containing epidermal  
 A:Reference number: A34412; MUID:89380326; PMID:2789221  
 A:Accession: A34412  
 A:Molecule type: protein  
 A:Residues: 85-126 <PER>  
 A:Note: beta-hydroxyaspartic acid site  
 R:Inoue, K.; Morita, T.  
 Eur. J. Biochem. 218, 153-163, 1993  
 A:Title: Identification of O-linked oligosaccharide chains in the activation peptides of  
 A:Reference number: S39414; MUID:94062825; PMID:8243461  
 A:Accession: S39414  
 A:Molecule type: protein  
 A:Residues: 183-196; 199-209; 216-233 <INO>  
 A:Note: carbohydrate binding sites  
 R:Titani, K.; Hermodson, M.A.; Fujikawa, K.; Ericsson, L.H.; Walsh, K.A.; Neurath, H.; D  
 Biochemistry 11, 4899-4903, 1972  
 A:Title: Bovine factor X-1a (activated Stuart factor). Evidence of homology with mammali  
 A:Reference number: A12453; MUID:73053314; PMID:4264286

A:Contents: annotation; active site  
 R:Fujikawa, K.; Titani, K.; Davie, E.W.  
 Proc. Natl. Acad. Sci. U.S.A. 72, 3359-3363, 1975  
 A:Title: Activation of bovine factor X (Stuart factor): conversion of factor Xaalpha  
 A:Reference number: A13504; MUID:76053121; PMID:1059122  
 A:Contents: annotation; activation  
 R:Sugo, T.; Bjork, I.; Holmgren, A.; Stenflo, J.  
 J. Biol. Chem. 259, 5705-5710, 1984  
 A:Title: Calcium-binding properties of bovine factor X lacking the gamma-carboxygluta  
 A:Reference number: A38024; MUID:84185716; PMID:6546930  
 A:Contents: annotation; calcium binding  
 R:Morita, T.; Jackson, C.M.  
 J. Biol. Chem. 261, 4008-4014, 1986  
 A:Reference number: A38025; MUID:86140210; PMID:3949800  
 A:Contents: annotation; sulfate binding  
 C:Comment: Factor Xa converts prothrombin to thrombin during blood clotting.  
 C:Comment: The two chains are formed from a single-chain precursor by the excision of  
 C:Comment: The activation peptide is cleaved by factor IXa (in the intrinsic pathway)  
 activation.  
 C:Comment: Calcium binds to the gamma-carboxyglutamic acid (Gla) residues and, with s  
 C:Comment: The gamma-carboxyglutamic acid residues arise by a posttranslational, vita  
 C:Gene: F10  
 A:Map position: 13q34  
 C:Function:  
 A:Description: catalyzes the proteolytic activation of prothrombin to thrombin in the  
 A:Pathway: blood coagulation  
 C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homol  
 C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglu  
 F:1-15/Domain: signal sequence #status predicted <SIG>  
 F:16-40/Domain: propeptide #status predicted <PRO>  
 F:25-84/Domain: Gla domain homology <GLA>  
 F:41-180/Product: coagulation factor X light chain #status experimental <LCH>  
 F:90-121/Domain: EGF homology <EG>  
 F:129-164/Domain: EGF homology <EG2>  
 F:183-492/Product: coagulation factor X heavy chain #status experimental <HCH>  
 F:183-233/Domain: activation peptide #status experimental <APT>  
 F:234-492/Product: coagulation factor Xa heavy chain #status experimental <AHC>  
 F:234-461/Domain: trypsin homology <TR>  
 F:46,47,54,56,59,60,65,66,69,72,75,79/Modified site: gamma-carboxyglutamic acid (Glu)  
 F:57-62,90-101,95-110,112-121,129-140,136-149,151-164,172-341/disulfide bonds: #statu  
 F:103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental  
 F:200/Binding site: sulfate (tyr) (covalent) (partial) #status experimental  
 F:208,485/Binding site: carbohydrate (thr) (covalent) #status experimental  
 F:218/Binding site: carbohydrate (asn) (covalent) #status experimental  
 F:233-234/Cleavage site: Arg-Ile (coagulation factor IXa, coagulation factor VIIa) #s  
 F:240-245,260-276,389-403,414-442/Disulfide bonds: #status experimental  
 F:275,321,418/Active site: His, Asp, Ser #status predicted  
 Alignment Scores:  
 Pred. No.: 8 46e-43 Length: 492  
 Score: 809.00 Matches: 171  
 Percent Similarity: 52.27% Conservative: 71  
 Best Local Similarity: 36.93% Mismatches: 149  
 Query Match: 34.62% Indels: 72  
 DB: 1 Gaps: 12  
 US-09-763-153-1 (1-1245) x EXBO (1-492)  
 QY 1 GCCAATCTCTTCTGGAGGAGCTCCGTCACAGAGCCTGGAGCGGAGTGCATAGAGGAG 60  
 Db 41 AlaAsnSerPheLeuGluValLysGlnGlyAsnLeuGluArgGluCysLeuGluGlu 60  
 QY 61 ATCTGTGACTTCGAGAGGCCAGGAAATTTCCAAATTTGATGATGACACTGGCCCTC 120  
 Db 61 AlaCysSerLeuGluAlaArgGluValPheGluAspAlaGluThrAspGluPhe 80  
 QY 121 TGTGTCAAGCACGTCGACGCTGACCGTGTGCTTGTCTTGGCTTGGAGCAGCCGTCGGCC 180  
 Db 81 TrpSerLysTyrLysAspGlyAspGlnCys-----GluGlyHisProCysLeu 96  
 QY 181 AGCTGTGTCTGCGGGACGCGCTGCATCGAGCGCATCGGACGCTTACAGCTCGGACTGC 240

Db 97 Asn -----GlnGlyHisCysLysAspGlylleGlyAspTyrThrCysThrCys 112  
QY 241 CGCAGCGCTGGAGGCGCTTCTGCCAG-----CGCAGGTGAGCTTCTCCTCAAT 291  
Db 113 AlaGluGlyPheGluGlyLysAsnCysGlnPheSerThrArgGluIle----- 128  
QY 292 TGCTCGCTGGACAGCGGGCTGCACCGATTACTGCTAGAGGAGTGGCTGGCGGCGC 351  
Db 129 CysSerLeuAspAsnGlyGlyCysAspGlnPheCysArgGluGluArgSerGluValArg 148  
QY 352 TGCTAGCTGCTGGCTGCTCAAGCTGGGGACGACCTCTGTCAGTGTACCCCGCGAGTG 411  
Db 149 CysSerCysAlaHisGlyTyrValLeuGlyAspAsnSerLysSerCysValSerThrGlu 168  
QY 412 AAGTTCCCTTGTGGAGG-----CCCTGGAAGCGGATGGAGAGAGCGACGCTCACTGAAA 468  
Db 169 ArgPheProCysGlyLysPheThrGlnGlyArgSerArgTyrAlaIleHisThrSer 188  
QY 469 CGAGCACAGAAGAC-----CAGAGACCAAGTAGATCG----- 504  
Db 189 GluAspAlaLeuAspAlaSerGluLeuGluHisTyrAspProAlaAspLeuSerProThr 208  
QY 504 ----- 504  
Db 209 GluSerSerLeuAspLeuLeuGlyLeuAsnArgThrGluProSerAlaGlyGluAspGly 228  
QY 505 -----CGGCTCATGATGGGAAGATGACCGCGGGGAGACAGCCCTGGCAG 552  
Db 229 SerGlnValValArgIleValGlyGlyArgAspCysAlaGluGlyGluCysProTyrGln 248  
QY 553 GTGGTCTGCTGGACTCAAGAGAGAGCTGCGCTGCGGGGAGTGTCTATCCACCCCTCC 612  
Db 249 AlaLeuLeuValAsnGluGluAsnGluGlyPheCysGlyGlyThrIleLeuAsnGluPhe 268  
QY 613 TGGTGTCTGACAGCGGCCCTGCTGATGATGATGATGATGATGATGATGATGATGATG 672  
Db 269 TyrValLeuThrAlaHisGlyLysHisGlnAlaLysArgPheThrValArgValGly 288  
QY 673 GAGTATGACCTCGCGCGCTGGGAGAGTGGAGCTGGACCTGGACATCAAGAGAGTCTTC 732  
Db 289 AspArgAsnThrGluGlnGluGluGlyAsnGluMetalHisGluValGluMetThrVal 308  
QY 733 GTCACCCCACTACAGCAAGAGACACACCAATGACATGACATGACATGACATGACATG 792  
Db 309 LysHisSerArgPheValLysGluThrTyrAspPheAspIleAlaValLeuArgLeuLys 328  
QY 793 CAGCCCGCCACCTCTCCGACACCATGCTGCTCCCTCCCGGACAGCGGCTTGCA 852  
Db 329 ThrProIleArgPheArgAsnValAlaProAlaCysLeuProGluLysAspTyrAla 348  
QY 853 GAGCCGAGCTCATCAGCGCGGCGGAGGAGC---CTCGTGAGGGGTGGGGTACAC 909  
Db 349 GluAlaThrLeu-----MetThrGlnLysThrGlyIleValSerGlyPheGly----- 364  
QY 910 AGCAGCCGAGAGAGAGGCGGCAAGAGAAACCGCACCTTCGCTCAACTTCATCAAGATT 969  
Db 365 -----ArgThrHisGluLysGlyArgLeuSerSerThrLeuLysMetLeuGluVal 381  
QY 970 CCCGTGTCGCCCAATGAGTGCAGCGAGGTGCATGACCAACATGGTCTGTGAGAACATG 1029  
Db 382 ProTyrValAspArgSerThrCysLysLeuSerSerPheThrIleThrProAsnMet 401  
QY 1030 CTGTGTCGGGCATCCTCGGGGACCGGAGGATGCTGCGAGGCGGACAGTGGGGGCGCC 1089  
Db 402 PheCysAlaGlyTyrAspThrGlnProGluAspAlaCysGlnGlyAspSerGlyGlyPro 421  
QY 1090 ATGGTGGCTCTCTCCAGCGCACCTGTTCTGCTGGGCTGTGTGAGTGGGGTGGAGGC 1149  
Db 422 HisValThrArgPheLysAspThrTyrPheValThrGlyIleValSerTyrGlyGly 441  
QY 1150 TGTGGGCTCTTCAACATCGGGTTTACACCAAGTACCGCGCTACCTGAGTGGATC 1209  
Db 442 CysAlaArgLysGlyLysPheGlyValTyrThrLysValSerAsnPheLeuLysTyrPile 461

QY 1210 -----CATGGGCACATCAGACACAG 1230  
Db 462 AspLysIleMetLysAlaArgAlaGlyAlaAlaGlySerArgGlyHis-----Ser 478  
QY 1231 GAAGCCCCC 1239  
Db 479 GluAlaPro 481  
RESULT 7  
EXCH  
coagulation factor Xa (EC 3.4.21.6) precursor - chicken  
N;Alternate names: virus-activating proteinase  
C;Species: Gallus gallus (chicken)  
C;Date: 12-Feb-1993 #sequence\_revision 07-Feb-1997 #text\_change 16-Jul-1999  
C;Accession: S15838; S20380; S20381  
R;Suzuki, H.; Harada, A.; Hayashi, Y.; Wada, K.; Asaka, J.; Gotoh, B.; Ogasawara, T.;  
FEBS Lett. 283, 281-285, 1991  
A;Title: Primary structure of the virus activating protease from chick embryo. Its id  
A;Reference number: S15838; MUID:91257322; PMID:2044767  
A;Accession: S15838  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-475 <SUZ>  
A;Cross-references: DDBJ:P00844; NID:g222869; PIDN:BAA00724.1; PID:g222870  
R;Gotoh, B.; Yamauchi, F.; Ogasawara, T.; Nagai, Y.  
FEBS Lett. 296, 274-278, 1992  
A;Title: Isolation of factor Xa from chick embryo as the amniotic endoprotease respon  
A;Reference number: S20380; MUID:92164779; PMID:1537403  
A;Accession: S20380  
A;Molecule type: protein  
A;Residues: 241-246, 'X', 248-251, 'X', 253-261 <GOT>  
C;Function:  
A;Description: catalyzes the proteolytic activation of prothrombin to thrombin in the  
A;Pathway: blood coagulation  
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homol  
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglu  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-40/Domain: propeptide #status predicted <PRO>  
F;25-84/Domain: Gla domain homology <GLA>  
F;81-185/Product: coagulation factor X light chain #status experimental <LCH>  
F;90-121/Domain: EGF homology <EG1>  
F;129-167/Domain: EGF homology <EG2>  
F;186-475/Product: coagulation factor X heavy chain #status predicted <HCH>  
F;241-475/Product: coagulation peptide #status predicted <APT>  
F;241-468/Domain: trypsin homology <TRY>  
F;46,47,54,56,59,60,65,66,69,72,75,79/Modified site: gamma-carboxyglutamic acid (Glu)  
F;57-62,90-101,95-110,112-121,129-140,136-152,154-167,175-348,247-252,267-283,396-410  
F;103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted  
F;196,207,228,285/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;282,328,425/Active site: His, Asp, Ser #status predicted

Alignment Scores:  
Pred. No.: 2,47e-42 Length: 475  
Score: 801.50 Matches: 163  
Percent Similarity: 54.77% Conservative: 84  
Best Local Similarity: 36.14% Mismatches: 147  
Query Match: 34.30% Indels: 57  
Gaps: 8

US-09-763-153-1 (1-1245) x EXCH (1-475)

QY 1 GCCAACTCTCTCTCGAGGAGCTCCCTCAGCAGCCTGGAGCGGAGTGCATAGAGGAG 60  
Db 41 AlaAsnSerPheLeuGluGluMetLysGlnGlyAsnGluArgGluCysAsnGluGlu 60  
QY 61 ATCTGTGACTTCGAGGAGCCCAAGAAATTTCCAAAATGTGGATGACACACTGGCCTTC 120  
Db 61 ArgCysSerLysGluGluAlaArgGluAlaPheGluAsnGluLysThrGluGluPhe 80

QY 121 TGGTCAAGACGCTGACGGTGTACCAAGTCTTGGTCTTCCCTTGGAGACACCGCGCGC 180  
Db 81 TTPAsnIleTyrValAspGlyAspGlnCysSerAsnPro----- 94  
QY 181 AGCCTGTGCTGGCGGACGACGAGTGTATCGACGGCATCGGAGCTTACAGTTCGCGACTGC 240  
Db 95 -----CysHisTyrGlyGlnCysLysAspGlyLeuGlySerTyrThrCysSerCys 112  
QY 241 CGCAGCGGCTGGAGGCGCTTCTCCAGCGGAGGTAGTCTTCAATTGCTCGGTG 300  
Db 113 LeuAspGlyTyrGlnGlyLysAsnCysGluPheValIlePro---LysTyrCysLysIle 131  
QY 301 GACAACGGCGGTGTACGACGATCTACTGCTAGAGAGGTGGCTGGCGGCGC----- 351  
Db 132 AsnAsnGlyAspCysGluGlnPheCysSerIleLysLysSerValGlnLysAspValVal 151  
QY 352 TGTAGTGTGCGCTGGCTTACAGCTGGGACGACCTCTCTGAGTGTACCCCGCAGTGT 411  
Db 152 CysSerCysThrSerGlyTyrGluLeuAlaGluAspGlyLysGlnCysValSerLysVal 171  
QY 412 AAGTTCCTCTGTGGAGGCGCTTGGAGCGGATGGAGAACGCGAGTCACTGAAACGA 471  
Db 172 LysTyrProCysGlyLysValLeuMetLysArgIleLysArgSerValIleLeuProThr 191  
QY 472 GAC-----ACAGAAGACCAA----- 486  
Db 192 AsnSerAsnThrAsnAlaThrSerAspGlnAspValProSerThrAsnGlySerIleLeu 211  
QY 486 ----- 486  
Db 212 GluGluValPheThrThrThrGluSerProThrProProArgAsnGlySerSer 231  
QY 487 -----GAAGACCAAGTAGATCCGCGCTCATTGTATGGAAGATACACAGCGGGGAC 540  
Db 232 IleThrAspProAsnValAspThrArgIleValGlyAspGluCysArgProGlyGlu 251  
QY 541 AGCCCTGTGACGGTGTCTGTGCTGCACTAAGAAAGCTGGCTGGCGGCGAGTGTCTC 600  
Db 252 CysProTrpGlnAlaValLeuIleAsnGluLysGlyGluPheCysGlyGlyThrIle 271  
QY 601 ATCCACCCCTCTCTGGTGTGACAGCGGCCCTCATGATGATGATGATCAAGAGCTCCT 660  
Db 272 LeuAsnGluAspPheIleThrAlaAlaHisCysIleAsnGlnSerLysGluLys 291  
QY 661 GTACGCTGTGAGATATGACCTGGCGCTGGGAGAGTGGAGCTGGACCTGGACATC 720  
Db 292 ValValValGlyGluValAspArgGluLysGluGluHisSerGluThrThrHisThrAla 311  
QY 721 AAGGAGGTCTGTGCTCCACCCCACTACAGCAAGACGACCCACCAATGATCGCACTG 780  
Db 312 GluLysIlePheValHisSerLysTyrIleAlaGluThrTyrAspAsnAspIleAlaLeu 331  
QY 781 CTGCACTGTGCGCCAGCCGACCTCTGCGAGACCATAGTGCCTGCTCCCGGAC 840  
Db 332 IleLysLeuLysGluProIleAlaPheSerGluTyrValValProAlaCysLeuProGln 351  
QY 841 AGCGGCTGTGAGAGCGGAGCTC---AATCAGCGCGGCGGACGAGACCTGTCACGCGC 897  
Db 352 AlaAspPheAlaAsnGluValLeuMetAsnGln-----LysSerGlyMetValSerGly 369  
QY 898 TGGGCTTACACAGCAGCCGAGAGAGGCGCCAGAGAACCCACCTTCTGCTCCTCAAC 957  
Db 370 PheGlyArgGluPheGluAlaGlyArgLeuSerLysArg-----LeuLys 384  
QY 958 TTCATCAAGATTCGGTGGTCCCGCATCATGATGTCAGCGAGGTATGAGCAACATGGT 1017  
Db 385 ValLeuGluValProTyrValAspArgSerThrCysLysGlnSerThrAsnPheAlaIle 404  
QY 1018 TCTGAGACATGCTGTGTGCGGATCTCTCGGACCGGACGAGGATGCTCGGAGGCGAC 1077  
Db 405 ThrGluAsnMetPheCysAlaGlyTyrGluThrGluLysAspAlaCysGlnGlyAsp 424

QY 1078 ACTGGGGGGCCCATGTGCTGCCTCCTCCACGCGACCTGTTCTCTGGTGGCGCTGTCGAGC 1137  
Db 425 SerGlyGlyProHisValThrArgTyrLysAspThrTyrPheValThrGlyLeuValSer 444  
QY 1138 TGGGTGTAGGCTGTGGCTCTCTCCACAACTACGCGCTTTACACCAAGTCAGCGCTAC 1197  
Db 445 TrpGlyGluGlyCysAlaArgLysGlyLysTyrGlyValTyrThrLysLeuSerArgPhe 464  
QY 1198 CTCGACTGGATCCATCGGCACATCAGACACAAG 1230  
Db 465 LeuArgTrpValArgThrValMetArgGlnLys 475  
RESULT 8  
EXHU  
coagulation factor Xa (EC 3.4.21.6) precursor [validated] - human  
N:Alternate names: Stuart factor  
C:Species: Homo sapiens (man)  
C:Date: 15-Nov-1984 #sequence.Revision 02-May-1994 #text\_change 08-Dec-2000  
C:Accession: A24478; JQ0917; A42485; A25853; A22208; A21284; A20362; S39415; I54051;  
R:Leytus, S.P.; Foster, D.C.; Kurachi, K.; Davie, E.W.  
Biochemistry 25, 5098-5102, 1986  
A:Title: Gene for human Factor X: a blood coagulation factor whose gene organization  
A:Reference number: A24478; MUID:87026600; PMID:3768336  
A:Accession: A24478  
A:Molecule type: DNA  
A:Residues: 1-488 <LEY>  
A:Cross-references: GB:L29433; GB:M14327; NID:9459809; PIDN:AAA52764.1; PID:g182831  
R:Messier, T.L.; Pittman, D.D.; Long, G.L.; Kaufman, R.J.; Church, W.R.  
Gene 99, 291-294, 1991  
A:Title: Cloning and expression in COS-1 cells of a full-length cDNA encoding human  
A:Reference number: JQ0917; MUID:91216473; PMID:1902434  
A:Accession: JQ0917  
A:Molecule type: mRNA  
A:Residues: 1-488 <MES>  
A:Cross-references: GB:M57285; NID:8182389; PIDN:AAA52421.1; PID:g182390  
R:Miao, C.H.; Leytus, S.P.; Chung, D.W.; Davie, E.W.  
J. Biol. Chem. 267, 7395-7401, 1992  
A:Title: Liver-specific expression of the gene coding for human factor X, a blood coa  
A:Reference number: A42485; MUID:92218390; PMID:1313796  
A:Accession: A42485  
A:Molecule type: DNA  
A:Residues: 1-15 <MIA>  
A:Experimental source: liver  
A:Note: sequence extracted from NCBI backbone (NCBIN:93780, NCBI:93787)  
R:Kaul, R.K.; Hildebrand, B.; Roberts, S.; Jagadeeswaran, P.  
Gene 41, 311-314, 1986  
A:Title: Isolation and characterization of human blood-coagulation factor X cDNA.  
A:Reference number: A25853; MUID:86221713; PMID:3011603  
A:Accession: A25853  
A:Molecule type: mRNA  
A:Residues: 19-284, 'E', 289-488 <KAU>  
A:Cross-references: GB:M22613; NID:gl80335; PIDN:AAA51984.1; PID:g180336  
R:Fung, M.R.; Hay, C.W.; MacGillivray, R.T.A.  
Proc. Natl. Acad. Sci. U.S.A. 82, 3591-3595, 1985  
A:Title: Characterization of an almost full-length cDNA coding for human blood coagul  
A:Reference number: A22208; MUID:85216545; PMID:2582420  
A:Accession: A22208  
A:Molecule type: mRNA  
A:Residues: 13-441, 'S', 443-488 <FUN>  
A:Cross-references: GB:K03194; NID:gl82840; PIDN:AAA52490.1; PID:g182841  
R:Leytus, S.P.; Chung, D.W.; Kistiel, W.; Kurachi, K.; Davie, E.W.  
Proc. Natl. Acad. Sci. U.S.A. 81, 3699-3702, 1984  
A:Title: Characterization of a cDNA coding for human factor X.  
A:Reference number: A21284; MUID:84222026; PMID:6587384  
A:Accession: A21284  
A:Molecule type: mRNA  
A:Residues: 13-284, 'E', 289-488 <LE2>  
A:Cross-references: GB:K01886  
R:McMullen, B.A.; Fujikawa, K.; Kistiel, W.; Sasagawa, T.; Howald, W.N.; Kwa, E.Y.; We  
Biochemistry 22, 2875-2884, 1983  
A:Title: Complete amino acid sequence of the light chain of human blood coagulation f  
A:Reference number: A20362; MUID:83252707; PMID:6871167  
A:Accession: A20362

A:Molecule type: protein  
A:Residues: 41-179 <MCM>  
R:Inoue, K.; Morita, T.  
Eur. J. Biochem. 218, 153-163, 1993  
A:Title: Identification of O-linked oligosaccharide chains in the activation peptides of  
A:Reference number: S39414; MUID:94062825; PMID:8243461  
A:Accession: S39415  
A:Molecule type: protein  
A:Residues: 183-234 <INO>  
A:Note: glycosylation sites  
A:Note: identification and characterization of beta-hydroxyaspartic acid  
R:Jagadeeswaran, P.; Reddy, S.V.; Rao, K.J.; Hamsabhusan, K.; Lyman, G.  
Gene 84, 517-519, 1989  
A:Title: Cloning and characterization of the 5' end (exon 1) of the gene encoding human  
A:Reference number: I54051; MUID:90128299; PMID:2612918  
A:Accession: I54051  
A>Status: translation not shown; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-23 <RES>  
A:Cross-references: GB:M33297; NID:g183860; PIDN:AAA52636.1; PID:g553330  
Padmanabhan, K.; Padmanabhan, K.P.; Tullinsky, A.; Park, C.H.; Bode, W.; Huber, R.; Bla  
Mol. Biol. 232, 947-966, 1993  
A:Title: Structure of human des(1-45) factor Xa at 2.2 angstroms resolution.  
Reference number: A49458; MUID:93360277; PMID:835279  
A:Contents: annotation; X-ray crystallography, 2.2 angstroms  
C:Comment: The two chains held together by one disulfide bond are formed from a single-c  
C:Comment: The activation peptide is cleaved by factor IXa (in the intrinsic pathway) or  
C:Genetics:  
A:Gene: GDB:F10  
A:Cross-references: GDB:119890; OMIM:227600  
A:Map position: 13q34-13q34  
A:Introns: 24/1; 77/3; 86/1; 124/1; 150/3; 249/3; 289/1  
A:Note: deficiency of this factor causes Stuart disease  
C:Function:  
A:Description: catalyzes the proteolytic activation of prothrombin to thrombin in the pr  
A:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology  
C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglut  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-40/Domain: propeptide #status predicted <PRO>  
F:25-84/Domain: Gla domain homology <GLA>  
F:41-179/Product: coagulation factor X light chain #status experimental <LCH>  
F:90-121/Domain: EGF homology <EGF>  
F:129-164/Domain: EGF homology <EG2>  
F:183-488/Product: coagulation factor X heavy chain #status experimental <HCH>  
F:183-234/Domain: activation peptide #status experimental <APT>  
F:235-488/Product: coagulation factor Xa heavy chain #status experimental <ACT>  
F:235-462/Domain: trypsin homology <TRY>  
F:46, 47, 54, 56, 59, 60, 65, 66, 69, 72, 79/Modified site: gamma-carboxyglutamic acid (Glu) #stat  
F:57-62/Disulfide bonds: #status predicted  
F:90-101, 95-110, 112-121, 129-140, 136-149, 151-164, 172-342, 241-246, 261-277, 390-404, 415-443/  
F:103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental  
F:199, 211/Binding site: carboxylate (Thr) (covalent) #status experimental  
F:21, 231/Binding site: carboxylate (Asn) (covalent) #status experimental  
F:34-235/Cleavage site: Arg-Ile (coagulation factor IXa, coagulation factor VIIa) #stat  
F:276, 322, 419/Active site: His, Asp, Ser #status experimental

Alignment Scores:  
Pred. No.: 2,47e-42 Length: 488  
Score: 801.50 Matches: 159  
Percent Similarity: 54.55% Conservative: 87  
Best Local Similarity: 35.25% Mismatches: 150  
Query Match: 34.30% Indels: 55  
DB: 1 Gaps: 8

US-09-763-153-1 (1-1245) x EXHU (1-488)

QY 1 GCCAACTCTCTCTGGAGGAGCTCCGTCACAGCAGCTGGAGGGAGTGTCATAGAGGAG 60  
|||||  
Db 41 AlaAsnSerPheLeuGluMetLysLysGlyHisLeuGluArgGluCysMetGluGlu 60  
|||||  
QY 61 ATCTGTGACATCGAGGAGGAGCAAGGAAATTTCCAAATGTGATGACACACTGGCTTC 120  
|||||  
Db 403 PheCysAlaGlyTyrAspThrLysGlnGluAspAlaCysGlnGlyAspSerGlyGlyPro 422

61 ThrCysSerTyrGluGluAlaArgGluValPheGluAspSerAspLysThrAsnGluPhe 80  
121 TGGTCCAAAGCACGTCGACGTCAGCTGCTTGGTCTTGGCCCTGGAGCACCCTGGCGC 180  
|||||  
Db 81 TrpAsnLysTyrLysAspGlyAspGlnCysGluThrSerPro----- 94  
181 AGCTGTGTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 240  
|||||  
Db 95 -----CysGlnAsnGlnGlyLysCysLysAspGlyLeuGlyGluTyrThrCysThrCys 112  
241 CGAGCGGCTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 300  
|||||  
Db 113 LeuGluGlyPheGluGlyLysAsnGlnGlyLysGluPheThrArgLysLeu---CysSerLeu 131  
301 GACAAAGCGGCTGTCAGCGCATTTACTGCTAGAGGAGGTGGGCTGGGCGGCGGCTGTAGCTGT 360  
|||||  
Db 132 AspAsnGlyAspCysAspGlnPheCysHisGluGluGlnAsnSerValValCysSerCys 151  
361 GCGCTGCTTACAAAGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420  
|||||  
Db 152 AlaArgGlyTyrThrLeuAlaAspAsnGlyLysAlaCysIleProThrGlyProTyrPro 171  
421 TGTGGGAGGCGGCTGGAGCGGATGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480  
|||||  
Db 172 CysGlyLys-----GlnThrLeuGluArgLysArgSerValAlaGlnAlaThrSer 189  
481 GACCAAGAGAGC----- 501  
190 SerSerGlyGluAlaProAspSerIleThrTriLysProTyrAspAlaAlaAspLeuAsp 209  
502 CCG----- 504  
210 ProThrGluAsnProPheAspLeuLeuAspPheAsnGlnThrGlnProGluArgGlyAsp 229  
505 -----CGCTCATTTGATGGGAAGATGACACCGGGGAGGAGGAGGAGGAGGAGGAGGAG 552  
230 AsnAsnLeuThrArgIleValGlyGlnGluCysLysAspGlyGluCysProTyrGln 249  
553 GTGTCTCTGTGACTCAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 612  
|||||  
Db 250 AlaLeuLeuIleAsnGluGluAsnGluGlyPheCysGlyGlyThrIleLeuSerGluPhe 269  
613 TGGGTGCTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 672  
|||||  
Db 270 TyrIleLeuThrAlaAlaHisCysLeuTyrGlnAlaLysArgPheLysValArgValGly 289  
673 GAGTATGACCTCGCGCGCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 732  
|||||  
Db 290 AspArgAsnThrGluGlnGluGlyGlyGlyAlaValHisGluValGluValValIle 309  
733 GTCCACCCCACTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 792  
|||||  
Db 310 LysHisAsnArgPheThrLysGluThrTyrAspPheAspIleAlaValLeuArgLeuLys 329  
793 CAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 852  
|||||  
Db 330 ThrProIleThrPheArgMetAsnValAlaProAlaCysLeuProGluArgAspTyrAla 349  
853 GAGCGCGAGCTCAATCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 909  
|||||  
Db 350 GluSerThrLeu-----MetThrGlnLysThrGlyIleValSerGlyPheGlyArgThr 367  
910 AGCAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 969  
|||||  
Db 368 HisGluLysGlyArgGlnSerThrArg-----LeuLysMetLeuGluVal 382  
970 CCGCTGGTCCGCAATGAGTGCAGCGGAGGTCATGACAAACATGGTGTCTGAGAAACATG 1029  
|||||  
Db 383 ProTyrValAspArgAsnSerCysLysLeuSerSerPheIleIleThrGlnAsnMet 402  
1030 CTGTGTGGCGGAGTCTCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1089  
|||||  
Db 403 PheCysAlaGlyTyrAspThrLysGlnGluAspAlaCysGlnGlyAspSerGlyGlyPro 422











Db 410 GlyHisValGlyValThrArgValSerArgAspThrGluTrpLeuSerArgLeuMet 429  
 QY 1222 AGAGACAAG 1230  
 Db 430 ArgSerLys 432  
 RESULT 12  
 A30351  
 coagulation factor IXa (EC 3.4.21.22) precursor - dog  
 C:Species: Canis lupus familiaris (dog)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A30351; I46201  
 R:Evans, J.P.; Watzke, H.H.; Ware, J.L.; Stafford, D.W.; High, K.A.  
 Blood 74, 207-212, 1989  
 A:Title: Molecular cloning of a cDNA encoding canine factor IX.  
 A:Reference number: A30351; MUID:89323338; PMID:2752110  
 A:Accession: A30351  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-452 <EVA>  
 A:Cross-references: GB:M21757; NID:g972719; PIDN:AAA75006.1; PID:g163948  
 Axelrod, J.H.; Read, M.S.; Brinkhous, K.M.; Verma, I.M.  
 Jc. Natl. Acad. Sci. U.S.A. 87, 5173-5177, 1990  
 Title: Phenotypic correction of factor IX deficiency in skin fibroblasts of hemophilic  
 A:Reference number: I46201; MUID:90311364; PMID:2367529  
 A:Accession: I46201  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-452 <AX>  
 A:Cross-references: GB:M33826; NID:g163949; PIDN:AAA30844.1; PID:g163950  
 C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology  
 C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutamic  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-40/Domain: propeptide #status predicted <PRO>  
 F:24-84/Domain: Gla domain homology <GLA>  
 F:41-452/Product: coagulation factor IX #status predicted <MAT>  
 F:90-121/Domain: EGF homology <EG1>  
 F:127-163/Domain: EGF homology <EG2>  
 F:218-445/Domain: trypsin homology <TRY>  
 F:46,47,54,56,59,60,65,66,69,72,75,79/Modified site: gamma-carboxyglutamic acid (Glu) #s  
 F:57-62,90-101,95-110,112-121,127-138,134-148,150-163,171-326,243-259,373-387,398-426/D  
 F:258,306,402/Active site: His, Asp, Ser #status predicted

Alignment Scores:  
 Pred. No.: 6,09e-40 Length: 452  
 Score: 763.00 Matches: 158  
 Percent Similarity: 53.97% Conservative: 73  
 Best Local Similarity: 36.92% Mismatches: 141  
 Query Match: 32.65% Indels: 56  
 DB: 1 Gaps: 12

us-09-763-153-1 (1-1245) x A30351 (1-452)  
 13 CTGAGAGAGCTCGCTCACAGCAGCTGGAGCGGAGTGCATAGAGAGATCTGTGACTTC 72  
 45 LeuGluGluPheValArgGlyAsnLeuGluArgGluCysIleGluGluLysCysSerPhe 64  
 73 GAGGAGGCCAAGGAATTTCCAAAATGTGGATGACACACTGGCCCTTCGTGTCACAGCAC 132  
 65 GluGluAlaArgGluValPheGluAsnThrGluLysThrThrGluPheTrpLysGlnTyr 84  
 133 GTCAGCGGTGACAGTGTGTGTCCTGGAGCACCCTGGAGCACCCTGGAGCACCCTGGAGC 192  
 85 ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu 96  
 193 GGCAGCGAGCTGCTGACGAGCCTCGGAGCTTCAGCTTCAGCTTCGAGCTGCGAGCGGCTG 252  
 97 AsnAspGlyValCysLysAspAspIleAsnSerTyrGluCysTrpCysArgAlaGlyPhe 116  
 253 GAGGCGCGCTCTGCCAGCGCGAGGTGAGCTTCCTCAATTCGTGTCGACACAGCGCGC 312  
 117 GluGlyLysAsnCysGluLeuAspValThr-----CysAsnIleLysAsnGlyArg 133

QY 313 TGCACGATTAAGTCTAGAGAGAGTGGCTGGCGCGCGC-----TGTAGCTGTGGC 363  
 Db 134 CysLysGlnPheCys-----LysLeuGlyProAspAsnLysValValCysSerCysThr 151  
 QY 364 CTGGGTACAAGCTGGGGGACGACCTCCTCGAGCTGACCCCGCAGTGAAGTTCCTTGT 423  
 Db 152 ThrGlyTyrGlnLeuAlaGluAspGlnArgSerCysGluProAlaValProPheProCys 171  
 QY 424 GGGAGG-----CCCTGGAAGCGATGGAGAAGAGCGGAGTCCATCGAAGACGAGAC 474  
 Db 172 GlyArgValSerValProHisIleSerMetThrArgThrArgAlaGluThrLeuPheSer 191  
 QY 475 ACAGAGACACCAAGACAGACCAAGTAGAT----- 501  
 Db 192 AsnMetAspTyrGluAsnSerThrGluValGluLysIleLeuAspAsnValThrGlnPro 211  
 QY 502 -----CCGCGGCTCATTGATGGAGAGTACACAGCGGGGAGACAGCCCTGG 549  
 Db 212 LeuAsnAspPheThrArgValValGlyGlyLysAspAlaLysProGlyGlnPheProTrp 231  
 QY 550 CAGGTGTCTGCTGCTGACTCAAAAGAAAGCTGGCTGGCGGCGAGTGTCTATCCACCCC 609  
 Db 232 Gln--ValLeuLeuAsnGlyLysValAspAlaPheCysGlySerIleIleAsnGlu 250  
 QY 610 TCCTGGGTGTGACAGCGGCCCACTGATGATGAGTCCCAAGAGCTCCTTGTTCAGGCT 669  
 Db 251 LysTrpValValThrAlaAlaHisCysIleGluProAspValLysIleThrIleValAla 270  
 QY 670 GGAGATGATCACCTGGCGGCTGGAGAGTGGAGCTGGACCTGGACATCAAGAGGTC 729  
 Db 271 GlyGluHisAsnThrGluLysArgGluHisThrGluGlnLysArgAsnValIleArgThr 290  
 QY 730 TTGCTCCACCCCACTACACAGAGACACACCC-----GACATGACATCGCAGCTGCTG 783  
 Db 291 IleLeuHisHisSerTyrAsnAlaThrIleAsnLysTyrAsnHisAspIleAlaLeuLeu 310  
 QY 784 CACCTGCCCGCCCGCCACCTCTCCAGACCATAGTCCCTCTCCCTCCCGGACAGC 843  
 Db 311 GluLeuAspGluProLeuThrLeuAsnSerTyrValThrProIleCys----- 326  
 QY 844 GGCCTTCAGAGCGCGAGTCAATCAG-----GCCGCCAGGAGAC 885  
 Db 327 ---IleAlaAspArgGluTyrSerAsnIlePheLeuLysPheGlySerGly----- 342  
 QY 886 CTCGTGAGCGGTGGGTACACAGCGCGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 945  
 Db 343 TyrValSerGlyTyrGlyArgValPheAsnLysGlyArgSerAlaSer----- 358  
 QY 946 TTCGTCTCAACTCATCAAGATTCCTGCTGCCGACCAATGAGTGCAGCGAGGTCA 1005  
 Db 359 ---IleLeuGlnTyrLeuLysValProLeuValAspArgAlaThrCysLeuArgSerThr 377  
 QY 1006 AGCAACATGTTGTGAGAACATCTGTGTGGGCGATCTCTGGGGGAGGAGGAGGAGGAG 1065  
 Db 378 LysPheThrIleTyrAsnAsnMetPheCysAlaGlyPheHisGluGlyLysAspSer 397  
 QY 1066 TCGAGGCGGACAGTGGGGGCGCCATGTCCTCTCCACGCGACCTGTTCTCTGTTG 1125  
 Db 398 CysGlnGlyAspSerGlyGlyProHisValThrGluValGluGlyIleSerPheLeuThr 417  
 QY 1126 GGCCTGTGTGAGTGGGTGAGGCGTGTGGGCTCCTTCCAACTACGCGGTTTACACCAA 1185  
 Db 418 GlyIleIleSerTrpGlyGluCysAlaMetLysGlyLysTyrGlyIleTyrThrLys 437  
 QY 1186 GTCAGCGCTACCTCGACTGGATC 1209  
 Db 438 ValSerArgTyrValAsnTrpIle 445  
 RESULT 13  
 KFHU  
 coagulation factor IXa (EC 3.4.21.22) precursor [validated] - human  
 N:Alternate names: antihemophilic factor B; Christinas factor  
 C:Species: Homo sapiens (man)

- C; Date: 17-Dec-1982 #sequence-revision 30-Jun-1987 #text change 15-Sep-2000  
 A; Accession: A00922; A37570; A30511; A32989; A22673; A21337; A37546; A30623; A60486; A20  
 R; Yoshitake, S.; Schach, B.G.; Postner, D.C.; Davie, E.W.; Kurachi, K.  
 Biochemistry 24, 3736-3750, 1985  
 A; Title: Nucleotide sequence of the gene for human factor IX (antithemophilic factor B).  
 A; Reference number: A00922; MUID:86000558; PMID:2994716  
 A; Accession: A00922  
 A; Molecule type: DNA  
 A; Residues: 1-461 <YOS>  
 A; Cross-references: GB:K02402; NID:g182612; PIDN:AA59620.1; PID:g182613  
 R; Anson, D.S.; Choo, K.H.; Rees, D.J.G.; Giannelli, F.; Gould, K.; Huddleston, J.A.; Bro  
 EMBO J. 3, 1053-1060, 1984  
 A; Title: The gene structure of human anti-haemophilic factor IX.  
 A; Reference number: A37570; MUID:84236100; PMID:6329734  
 A; Accession: A37570  
 A; Molecule type: DNA  
 A; Residues: 1-461 <ANS>  
 A; Cross-references: GB:K02048  
 R; Reitsma, P.H.; Bertina, R.M.; Ploos van Amstel, J.K.; Riemsens, A.; Briet, E.  
 Blood 72, 1074-1076, 1988  
 A; Title: The putative factor IX gene promoter in hemophilia B Leyden.  
 A; Reference number: A30511; MUID:88327116; PMID:3416069  
 A; Accession: A30511  
 A; Molecule type: DNA  
 A; Residues: 8-24 <REI>  
 A; Cross-references: EMBL:X55008; NID:g311288; PIDN:CA838245.2; PID:g4469253  
 R; Koerber, D.D.; Bottema, C.D.K.; Buerstedde, J.M.; Sommer, S.S.  
 Am. J. Hum. Genet. 45, 448-457, 1989  
 A; Title: Functionally important regions of the factor IX gene have a low rate of polymor  
 A; Reference number: A32989; MUID:89371752; PMID:2773937  
 A; Accession: A32989  
 A; Status: not compared with conceptual translation  
 A; Molecule type: DNA  
 A; Residues: 30-92 <KOE>  
 R; McGraw, R.A.; Davis, L.M.; Noyes, C.M.; Lundblad, R.L.; Roberts, H.R.; Graham, J.B.; S  
 Proc. Natl. Acad. Sci. U.S.A. 82, 2847-2851, 1985  
 A; Title: Evidence for a prevalent dimorphism in the activation peptide of human coagulat  
 A; Reference number: A22673; MUID:85190593; PMID:3857619  
 A; Accession: A22673  
 A; Molecule type: mRNA  
 A; Residues: 1-193, 'T', 195-461 <MCG>  
 A; Cross-references: GB:M1309; NID:g180552; PIDN:AAA52023.1; PID:g180553  
 A; Note: The authors translated the codon ACA for residue 29 as Tyr  
 R; Jaye, M.; de la Salle, H.; Schamber, F.; Balland, A.; Kohli, V.; Findeli, A.; Tolstosh  
 Nucleic Acids Res. 11, 2325-2335, 1983  
 A; Title: Isolation of a human anti-haemophilic factor IX cDNA clone using a unique 52-ba  
 A; Reference number: A21337; MUID:83220788; PMID:6687940  
 A; Accession: A21337  
 A; Molecule type: mRNA  
 A; Residues: 1-193, 'T', 195-461 <JAY>  
 A; Cross-references: GB:J00137; NID:g182610; PIDN:AAA52763.1; PID:g182611  
 R; Jagadeeswaran, P.; Lavellie, D.E.; Kaul, R.; Mohandas, T.; Warren, S.T.  
 Somat. Cell Mol. Genet. 10, 465-473, 1984  
 A; Title: Isolation and characterization of human factor IX cDNA: identification of Tag I  
 A; Reference number: A37546; MUID:84300526; PMID:6089357  
 A; Accession: A37546  
 A; Molecule type: mRNA  
 A; Residues: 38-193, 'T', 195-326 <JAG>  
 A; Cross-references: GB:M35672  
 R; Kurachi, K.; Davie, E.W.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 6461-6464, 1982  
 A; Title: Isolation and characterization of a cDNA coding for human factor IX.  
 A; Reference number: A30623; MUID:83065193; PMID:6959130  
 A; Accession: A30623  
 A; Molecule type: mRNA  
 A; Residues: 1-12, 'S', 'G', '84-203, 'P', 205-216, 'G', 218-298, 'A', 299-356, 'A',  
 A; Cross-references: GB:J00136; NID:g182608; PIDN:AAA98736.1; PID:g182609  
 R; Tharakan, J.; Strickland, D.; Burgess, W.; Drohan, W.N.; Clark, D.B.  
 Vox Sang. 58, 21-29, 1990  
 A; Title: Development of an immunoaffinity process for factor IX purification.  
 A; Reference number: A60486; MUID:90194857; PMID:2316207  
 A; Accession: A60486
- A; Molecule type: protein  
 A; Residues: 47-52, 'XX', 55-60, 'X', 62, 'XX', 65 <THA>  
 R; McMullen, B.A.; Fujikawa, K.; Kisiel, W.  
 Biochem. Biophys. Res. Commun. 115, 8-14, 1983  
 A; Title: The occurrence of beta-hydroxyaspartic acid in the vitamin K-dependent blood  
 A; Reference number: A20274; MUID:83308813; PMID:6688526  
 A; Accession: A20274  
 A; Molecule type: protein  
 A; Residues: 105-109, 'X', 111-115 <MCM>  
 R; Balland, A.; Faure, T.; Carvallo, D.; Cordier, P.; Ulrich, P.; Fournet, B.; de la S  
 Eur. J. Biochem. 172, 565-572, 1988  
 A; Title: Characterisation of two differently processed forms of human recombinant fac  
 A; Reference number: S02527; MUID:88166735; PMID:3280312  
 A; Accession: S02527  
 A; Molecule type: protein  
 A; Residues: 29-63 <BAL>  
 R; Jallat, S.; Perraud, F.; Dalemans, W.; Balland, A.; Dieterle, A.; Faure, T.; Meulle  
 EMBO J. 9, 3295-3301, 1990  
 A; Title: Characterization of recombinant human Factor IX expressed in transgenic mice  
 A; Reference number: S12058; MUID:91006024; PMID:2209546  
 A; Accession: S12058  
 A; Molecule type: mRNA; protein  
 A; Residues: 1-68 <JAL>  
 A; Note: processed forms expressed in recombinant system  
 R; Handford, P.A.; Baron, M.; Mayhew, M.; Willis, A.; Beesley, T.; Brownlee, G.G.; Cam  
 EMBO J. 9, 475-480, 1990  
 A; Title: The first EGF-like domain from human factor IX contains a high-affinity calc  
 A; Reference number: S12377; MUID:90151623; PMID:2406129  
 A; Accession: S12377  
 A; Molecule type: protein  
 A; Residues: 92-130 <HAN>  
 A; Note: NMR detection of calcium binding by domain expressed in recombinant system  
 R; de la Salle, C.; Charmantier, J.L.; Baas, M.J.; Schwartz, A.; Wiesel, M.L.; Grunbe  
 Thromb. Haemost. 70, 370-371, 1993  
 A; Title: A deletion located in the 3' non translated part of the factor IX gene respo  
 A; Reference number: I59612; MUID:94054330; PMID:8236150  
 A; Accession: I59612  
 A; Status: translated from GB/EMBL/DDBJ  
 A; Molecule type: DNA  
 A; Residues: 444-461 <RES>  
 A; Cross-references: GB:S66752; NID:g439773; PIDN:AAB28588.1; PID:g439774  
 R; Stoflet, E.S.; Koerber, D.D.; Sarkar, G.; Sommer, S.S.  
 Science 239, 491-494, 1988  
 A; Title: Genomic amplification with transcript sequencing.  
 A; Reference number: I59529; MUID:88127096; PMID:3340835  
 A; Accession: I59529  
 A; Status: translated from GB/EMBL/DDBJ  
 A; Molecule type: DNA  
 A; Residues: 290-359 <RE2>  
 A; Cross-references: GB:M19063; NID:g182622; PIDN:AAA52456.1; PID:g182623  
 R; Agarwala, K.L.; Kawabata, S.; Takao, T.; Murata, H.; Shimonishi, Y.; Nishimura, H.;  
 Biochemistry 33, 5167-5171, 1994  
 A; Title: Activation peptide of human factor IX has oligosaccharides O-glycosidically  
 A; Reference number: A54255; MUID:94227047; PMID:8172892  
 A; Accession: A54255  
 A; Molecule type: protein  
 A; Residues: 'D', 204, 'X', 206-211, 212, 'D', 214, 'X', 216-221, 'D' <AGA>  
 A; Note: the residues designated 'X' were determined to be threonine bound to carbohy  
 R; Di Scipio, R.G.; Kurachi, K.; Davie, E.W.  
 J. Clin. Invest. 61, 1528-1538, 1978  
 A; Title: Activation of human factor IX (Christmas factor).  
 A; Reference number: A18483; MUID:78194509; PMID:659613  
 R; McGraw, R.A.; Davis, L.M.; Noyes, C.M.; Graham, J.B.; Roberts, H.R.; Stafford, D.W.  
 Am. Soc. Hematol. Abstr. 64(Suppl.1), 262a, 1984  
 A; Reference number: A37569  
 A; Contents: annotation  
 A; Note: 194-Thr was also found  
 R; Morita, T.; Isaacs, B.S.; Esmon, C.T.; Johnson, A.E.  
 J. Biol. Chem. 259, 5698-5704, 1984  
 A; Title: Derivatives of blood coagulation factor IX contain a high affinity Ca2+-bind  
 A; Reference number: A37543; MUID:84185715; PMID:6425296

US-09-763-153-1 (1-1245) x KFHU (1-461)

	QY	13	CTGGAGGAGCTCCGTACACAGCAGCCTGGAGCGGGAGTCATAGAGGAGATCTGTCACTTC	72
	Dbb	52	LeuGluGluPheValIcInGlyAsnLeuGluArgGluCysMetGluGluLysCysSerPhe	71
	QY	73	GAGGAGGCCAAGAATAATTCCAAATGTGGATGACACACTGGCCTTCTGTCTCAAGCAC	132
	Dbb	72	GluGluAlaArgGluValPheGluAsnThrGluArgThrThrGluPheTrpLysGlnItyr	91
	QY	133	GTCCAGCGTGACCAGTGCCTGTTCGTCCTTTGGAGACCCCCTGGCCGACCCCTGTGCTGCC	192
	pB	92	ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu	103

QY	1018	TCTGAGAACATGCTGTGTGCGGGCATCTCTGGGGACCGCAGGATGCCTCTCGAGGGCGAC	107
DB	391	TyrAsnAsnMetPheCysAlaGlyPheHisGluClyArgaspSerCysGlnGlyAsp	410
QY	1078	ACTGGGGGGCCCATGCTCGCCTCTCTCACGGCACCTGGTTCCTGTGGCTGGCGCTGTGAGC	113
DB	411	SerGlyGlyProHisValThrGluValGluClyThrSerPheLeuThrGlyIleIleSer	430
QY	1138	TGGGTGAGGGCTGTGGCTCTCTCACAACTACGGCGCTTACACAAAGTCAGCGCGTAC	119
DB	431	TrpGlyGluGluCysAlaMetLysGlyLysTyrGlyIleYrThrLysValSerArgTyr	450

QY 1198 CTGCGACTGGATC 1209  
 Db 451 ValAsnTrpIle 454  
 RESULT 14  
 JQ0419  
 coagulation factor IXa (EC 3.4.21.22) precursor - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 16-Jul-1999  
 C:Accession: JQ0419; 149667  
 R:Wu, S.W.; Stafford, D.W.; Ware, J.  
 Gene 86, 275-278, 1990  
 A:Title: Deduced amino acid sequence of mouse blood-coagulation factor IX.  
 A:Reference number: JQ0419; MUID:90215309; PMID:2323576  
 A:Accession: JQ0419  
 A:Reference number: 146580; MUID:90152675; PMID:2303254  
 A:Accession: 149667  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 168-362, 'Q', 364-387, 'I', 389-451 <RES>  
 A:Cross-references: GB:M23109; NID:g193317; PIDN:AAA37629.1; PID:g387158  
 R:Sarkar, G.; Koebler, D.D.; Sommer, S.S.  
 Genomics 6, 133-143, 1990  
 A:Title: Direct sequencing of the activation peptide and the catalytic domain of the factor IX.  
 A:Reference number: 146580; MUID:90152675; PMID:2303254  
 A:Accession: 149667  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 168-362, 'Q', 364-387, 'I', 389-451 <RES>  
 A:Cross-references: GB:M26236; NID:g193319; PIDN:AAA37630.1; PID:g193320  
 C:Comment: This protein plays a critical role in blood coagulation.  
 C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology  
 C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutamate  
 F:1-16/Domain: signal peptide (fragment) #status predicted <SIG>  
 F:17-34/Domain: signal peptide (fragment) #status predicted <PRO>  
 F:19-79/Domain: Gla domain homology <GLA>  
 F:35-459/Product: coagulation factor IX #status predicted <MAT>  
 F:85-116/Domain: EGF homology <EGF>  
 F:122-158/Domain: EGF homology <EG2>  
 F:225-452/Domain: trypsin homology <TRY>  
 F:41-42,49,51,54,55,60,61,64,67,70,74/Modified site: gamma-carboxyglutamic acid (Glu) #  
 F:52-57,85-96,90-105,107-116,122-133,129-143,145-158,166-333,250-266,380-394,405-433/Dis  
 F:265,313,409/Active site: His, Asp, Ser #status predicted

Alignment Scores:  
 Pred. No.: 1.22e-37  
 Score: 726.00  
 Length: 459  
 Matches: 155  
 Percent Similarity: 50.91%  
 Conservative: 68  
 Mismatches: 151  
 Indels: 64  
 Gaps: 12

US-09-763-153-1 (1-1245) x JQ0419 (1-459)

QY 13 CTGAGGAGTCCGTCACAGCAGCGCTGGAGGAGTCATAGAGAGATCTGTGACTTC 72  
 Db 40 LeuGluGluPheValArgGlyAsnLeuGluGluCysIleGluGluArgCysSerPhe 59  
 QY 73 GAGAGGCGAAGAAATTTCCAAATGTGGATGACACTGCGCTTCTGGTCCCAAGCAC 132  
 Db 60 GluGluAlaArgGluValPheGluAsnThrGluLysThrThrGluPheTrpLysGlnTyr 79  
 QY 133 GTCCAGCGTGACCAAGTCTGTGCTTCCCTTGGAGCACCCGCTGGCGGACCTGTGCTGC 192  
 Db 80 ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu 91  
 QY 193 GGGCAGCGCAGTCATCGACGCGCTCGGAGCTTCAGCTGCGACTGCGCAGCGCTGG 252  
 Db 92 AsnGlyGlyLeuCysLysAspAspIleSerTyrGluCysTrpCysGlnValGlyPhe 111  
 QY 253 GAGGCGCGCTTCTGCCAGCGGAGGTGACCTTCAATTCGCTCGCTGACCAACGCGCG 312  
 Db 112 GluGlyArgAsnGluLeuAspAlaThr-----CysAsnIleLysAsnGlyArg 128  
 QY 313 TGCACGCAATTACTGCCTAGAGGAGTGGCTGCGCG---CGCTGAGCTGTGCGCGCTGCG 369

Db 129 CysLysGlnPheCysLysAsnSerProAspAsnLysValIleCysSerCysThrGluGly 148  
 QY 370 TACAAGCTGGGGAGGACGACCTCTCGAGTGTACCCCGAGTGAAGTTCCTTGTGGAGG 429  
 Db 149 TyrGlnLeuAlaGluAspGlnLysSerCysGluProThrValProPheProCysGlyArg 168  
 QY 430 CCC-----TGGAGCGGATG 444  
 Db 169 AlaSerIleSerTyrSerSerLysLysIleThrArgAlaGluThrValPheSerAsnMet 188  
 QY 445 GAGAAGAGCGCAGCTAC----- 462  
 Db 189 AspTyrGluAsnSerThrGluAlaValPheIleGlnAspAspIleThrAspGlyAlaIle 208  
 QY 463 CTGAACAGCAGACAGAACAGACCAAGACCAAGTAGAT---CCGCGCTCATGTGGG 519  
 Db 209 LeuAsnAsnValThrGluSerSerGluSerLeuAsnAspPheThrArgValValGlyGly 228  
 QY 520 AAGATGACCGGGGGGAGAGACGCCCTGGCAGGTGTCTGTGGTACTCAAAAGAAAG 579  
 Db 229 GluAsnAlaLysProGlyGlnIleProTrpGlnValIle---LeuAsnGlyGluIleGlu 247  
 QY 580 CTGSCCTGGCGGCGAGTGTCTATCCACCCCTCTGGTGTGTGACAGCGGCCCTCATG 639  
 Db 248 AlaPheCysGlyAlaIleIleAsnGluLysTrpIleValThrAlaAlaHisCysLeu 267  
 QY 640 GATGAGTCCAAGAGCTCTCTGTCAGGCTTGGAGAGTATGACCTGGCGGCTGGAGAG 699  
 Db 268 LysProGlyAspLysIleGluValAlaGlyGluThrAsnIleAspLysLysGluAsp 287  
 QY 700 TGGGAGCTGACCTGGACATCAGGAGGTCTTCTGTCACCCCACTACACAGACACC 759  
 Db 288 ThrGluGlnArgArgAsnValIleArgThrIleProHisGlnTyrAsnAlaThrIle 307  
 QY 760 ACC-----GACAAATGACATCGCAGCTGTGCACCTGGCGGCGGCTGGAGAGC 813  
 Db 308 AsnLysTyrSerHisAspIleAlaLeuLeuLysLysProLeuIleLeuAsnSer 327  
 QY 814 ACCATAGTCCCTATCTGCTCCCGGACAGCGGCTTGCAGAGCGGAGCTCAATCAG--- 870  
 Db 328 TyrValThrProIleCys-----ValAlaAsnArgGluTyrThrAsnIle 342  
 QY 871 -----CCGCGCAGGAGACCTCTGTCAGGCTGGGGCTGGGCTTACACAGACG 915  
 Db 343 PheLeuLysPheGlySerGly-----TyrValSerGlyTrpGlyLysValPheAsn 359  
 QY 916 CGAGAGAAGGAGGCCAAGAAACCGCACCTTCTGCTCAACTTCATCAAGATTCCCGTG 975  
 Db 360 LysGlyArgHisAlaSer-----IleLeuGlnTyrIleuArgValProLeu 374  
 QY 976 GTCCCGCCCAATGAGTGCAGGAGGTATGAGCAACATGGTGTCTGAGAACATGTGTGT 1035  
 Db 375 ValAspArgAlaThrCysLeuArgSerThrPheThrThrTyrAsnAsnMetPheCys 394  
 QY 1036 GCGGGCATCTCGGGACCGGAGGATGCTCGGAGGCGGACAGTGGGGGCGCCATGGTC 1095  
 Db 395 AlaGlyTyrArgGluGlyGlyLysAspSerCysGluGlyAspSerGlyGlyProHisVal 414  
 QY 1096 GCCTCCTCCAGCGGACCTGCTGCTGGGGCTGGTGTGAGTGGGTGGGGCTGGGTGGG 1155  
 Db 415 ThrGluValGluGlyThrSerPheLeuThrGlyIleIleSerTrpGlyGluGluCysAla 434  
 QY 1156 CTCCTTCAACATAGCGGCTTTACACCAAGTCAACCGCTACCTGACTGGATC 1209  
 Db 435 MetLysGlyLysTyrGlyIleTyrThrLysValSerArgTyrTyrValAsnTrpIle 452  
 RESULT 15  
 KFO  
 coagulation factor IXa (EC 3.4.21.22) precursor - bovine  
 N:Alternate names: Christmas factor  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 30-Nov-1980 #sequence\_revision 03-Aug-1984 #text\_change 16-Jul-1999

C;Accession: A14757; B20274; I45891; A00923  
R;Katayama, K.; Ericsson, L.H.; Enfield, D.L.; Walsh, K.A.; Neurath, H.; Davie, E.W.; T.  
Proc. Natl. Acad. Sci. U.S.A. 76, 4990-4994, 1979  
A;Title: Comparison of amino acid sequence of bovine coagulation factor IX (Christmas fa  
A;Reference number: A14757; MUID:80056619; PMID:291916  
A;Accession: A14757  
A;Molecule type: protein  
A;Residues: 1-63, 'T', 65-416 <KAT>  
R;McMullen, B.A.; Fujikawa, K.; Kistel, W.  
Biochem. Biophys. Res. Commun. 115, 8-14, 1983  
A;Title: The occurrence of beta-hydroxyaspartic acid in the vitamin K-dependent blood co  
A;Reference number: A20274; MUID:83308813; PMID:6688526  
A;Accession: B20274  
A;Molecule type: protein  
A;Residues: 59-63, 'X', 65-69 <MCM>  
R;Choo, K.H.; Gould, K.G.; Rees, D.J.G.; Brownlee, G.G.  
Nature 299, 178-180, 1982  
A;Title: Molecular cloning of the gene for human anti-haemophilic factor IX.  
A;Reference number: I45891; MUID:82272386; PMID:6287289  
A;Accession: I45891  
A;Title: Status: translated from GB/EMBL/DDJB  
A;Molecule type: mRNA  
A;Residues: 52-139 <CHO>  
Cross-references: GB:J00007; NID:gi163053; PIDN:AAA30520.1; PID:gi163054  
R;Hase, S.; Kawabata, S.; Nishimura, H.; Takeya, H.; Sueyoshi, T.; Miyata, T.; Iwanaga,  
J. Biochem. 104, 867-868, 1988  
A;Title: A new trisaccharide sugar chain linked to a serine residue in bovine blood coag  
A;Reference number: A44556; MUID:89213999; PMID:3149637  
A;Contents: annotation  
A;Note: structure and location of a carbohydrate covalently bound to Ser  
C;Comment: Factor IX is activated by factor X, which excises the activation peptide ph  
C;Comment: The gamma-carboxyglutamic acid residues arise by posttranslational, vitamin K  
C;Comment: Calcium binds to the gamma-carboxyglutamic acid (Gla) residues and, with str  
C;Function:  
A;Description: catalyzes the proteolytic activation of coagulation factor X in the prese  
A;Pathway: blood coagulation intrinsic pathway  
C;Superfamily: coagulation factor X; BGF homology; Gla domain homology; trypsin homology  
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam  
F.1-146/Product: coagulation factor IXa light chain #status experimental <ALC>  
F.1-45/Domain: Gla domain homology (fragment) <GLA>  
F.51-82/Domain: EGF homology <EGF>  
F.88-124/Domain: EGF homology <EG2>  
F.147-181/Domain: activation peptide #status experimental <APT>  
F.182-416/Product: coagulation factor IXa heavy chain #status experimental <AHC>  
F.182-409/Domain: trypsin homology <TRY>  
F.7, 8, 15, 17, 20, 21, 26, 27, 30, 33, 36, 40/Modified site: gamma-carboxyglutamic acid (Glu) #sta  
F.18-23, 51-62, 56-71, 73-82, 88-99, 95-109, 111-124, 132-290, 207-223, 337-351, 362-390/Disulfide  
F.53/Binding site: carboxylate (Ser) (covalent) #status experimental  
F.64/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental  
F.158, 168, 173, 261/Binding site: carboxylate (Asn) (covalent) #status experimental  
F.222, 270, 366/Active site: His, Asp, Ser #status predicted

Alignment Scores:  
Seq. No.: 6, 27e-37 Length: 416  
Seq. 1: 714.50 Matches: 148  
Percent Similarity: 51.64% Conservative: 72  
Best Local Similarity: 34.74% Mismatches: 157  
Query Match: 30.57% Indels: 49  
DB: 1 Gaps: 10

US-09-763-153-1 (1-1245) x KFB0 (1-416)  
QY 13 CTGGAGAGCTCCGTCACAGCAGCTGGAGCGGGAGTGCATAGAGAGATCTGTGACTTC 72  
Db 6 LeuGluGluPheValArgLysAsnLeuGluArgGluCysLysGluGluLysCysSerPhe 25  
QY 73 GAGGAGCCAGGAATTTCCAAATGTGGATGACACACTGGCTTCTGGTCCCAAGAC 132  
Db 26 GluGluAlaArgGluValPheGluAsnThrGluLysThrThrGluPheThrLysGlnTyr 45  
QY 133 GTGACGGTGGACAGTGTCTGTTCCTGGAGTGGAGCACCCTGGCCAGCGCTGCTGTC 192  
Db 46 ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu 57

QY 193 GGGCAGCGCAGCTGCATGACGGGATCGGAGCTTACGCTGCGCTGCGCGCGGCTGG 252  
Db 58 AsnGlyMetCysLysAspAspAspAspAspAspAspAspAspAspAspAspAsp 77  
QY 253 GAGGCGCGCTTCTGCAGCGGAGGTGAGCTTCTCTCAATTGCTCGCTGAGCAACAGCGGCG 312  
Db 78 GluGlyThrAsnCysGluLeuAspAlaThr-----CysSerIleLysAsnGlyArg 94  
QY 313 TGCACGCATTTACTGCTAGAGAGTGGCTGGCGG---CGCTGTAGCTGTCGCGCTGGC 369  
Db 95 CysLysGlnPheCysLysArgAspThrAspAsnLysValValCysSerCysThrAspGly 114  
QY 370 TACAAGCTGGGGAGCAGCTCTCCAGTGTCAACCCGAGTGAAGTGAAGTTCCTTGTGGGAGG 429  
Db 115 TyrArgLeuAlaGluAspGlnLysSerCysGluProAlaValProPheProCysGlyArg 134  
QY 430 CCC-----TGGAGCGGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 474  
Db 135 ValSerValSerHisIleSerLysLysLeuThrArgAlaGluThrIlePheSerAsnThr 154  
QY 475 ACAGAGACCAAGAGACCAAGTAGAT----- 501  
Db 155 AsnTyrGluAsnSerSerGluAlaGluIleIleIleIleIleIleIleIleIleIleIle 174  
QY 502 -----CCGCGGCTCATTTGATGGAGAGATGACAGCGGGGAGAGAGAGAGAGAG 546  
Db 175 SerPheAspGluPheSerArgValValGlyGlyGluAspAlaGluArgGlyGlnPhePro 194  
QY 547 TGCAGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 606  
Db 195 TrpGln---ValLeuLeuHisGlyGluIleAlaAlaPheCysGlyGlySerIleValAsn 213  
QY 607 CCCTCTGGTGTGACAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 666  
Db 214 GluLysTrpValValThrAlaAlaHisCysIleLysProGlyValLysIleThrValVal 233  
QY 667 CTGGAGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 726  
Db 234 AlaGlyGluHisAsnThrGluLysProGluProThrGluGlnLysArgAsnValIleArg 253  
QY 727 GTCTTCTGCCACCCCACTACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
Db 254 AlaIleProTyrHisSerTyrAsnAlaSerIleAsnLysTyrSerHisAspIleAlaLeu 273  
QY 781 CTGACCTGCCCGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 274 LeuGluLeuAspGluProLeuGluLeuAsnSerTyrValThrProIleCys----- 290  
QY 841 AGCGGCTTGCAGAGCGGAGCTCAATCAGCGCGGCGGAG-----ACCCCTCGTG 891  
Db 291 -----IleAlaAspArgAspTyrThrAsnIlePheSerLysPheGlyTyrGlyTyrVal 308  
QY 892 ACGGCTGGGGCTACACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 951  
Db 309 SerGlyTrpGlyLysValPheAsnArgGlyArgSerAlaSer-----Ile 323  
QY 952 CTCAACTTCATCAAGATCCCGTGGTCCCGCACAATGAGTGCAGCGAGGTCATGAGCAAC 1011  
Db 324 LeuGlnTyrLeuLysValProLeuValAspArgAlaThrCysLeuArgSerThrLysPhe 343  
QY 1012 ATGTGCTGTGAGAACATGCTGTGCTGCGGCGATCCTCGGGAGCGGAGGATGCTGCGAG 1071  
Db 344 SerIleTyrSerHisMetPheCysAlaGlyTyrHisGluGlyLysAspSerCysGln 363  
QY 1072 GGCACAGTGGGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1131  
Db 364 GlyAspSerGlyGlyProHisValThrGluValGluGlyThrSerPheLeuThrGlyIle 383  
QY 1132 GTGAGCTGGGTGAGGCTGTGGGCTCCTTCAACAACACAGCTGCTTACACCAAGTCAAG 1191  
Db 384 IleSerTrpGlyGluGluCysAlaMetLysGlyLysTyrGlyIleTyrThrLysValSer 403

QY 1192 CGCTACCTCGACTGGATC 1209

|||||:|||||

Db 404 ArgTyrValAsnTrpIle 409

RESULT 16

TBHU

thrombin [EC 3.4.21.5] precursor [validated] - human  
N:Alternate names: coagulation factor II

N:Contains: prothrombin

C:Species: Homo sapiens (man)

C:Date: 30-Nov-1980 #sequence\_revision 22-Jul-1994 #text\_change 08-Dec-2000

C:Accession: A29351; A00914; B00914; A37549; A37550; I51952

R:Degen, S.J.F.; Davie, E.W.

Biochemistry 26, 6165-6177, 1987

Title: Nucleotide sequence of the gene for human prothrombin.

Reference number: A29351; MUID:88077877; PMID:2825773

Accession: A29351

Molecule type: DNA

A:Residues: 1-622 &lt;DEG&gt;

A:CROSS-references: GB:M33691; NID:g558069; PIDN:AAC63054.1; PID:g339641

R:Degen, S.J.F.; MacGillivray, R.T.A.; Davie, E.W.

Biochemistry 22, 2087-2097, 1983

A:Title: Characterization of the complementary deoxyribonucleic acid and gene coding for

A:Reference number: A00914; MUID:83231469; PMID:6305407

A:Accession: A00914

A:Molecule type: mRNA

A:Residues: 8-163, 'N', 165-622 &lt;DE2&gt;

A:CROSS-references: GB:V00595; GB:J00307; NID:g37128; PIDN:CAA23842.1; PID:g1335344

A:Accession: B00914

A:Molecule type: DNA

A:Residues: 188-311 &lt;DE3&gt;

R:Walz, D.A.; Hewett-Emmett, D.; Seegers, W.H.

Proc. Natl. Acad. Sci. U.S.A. 74, 1969-1972, 1977

A:Reference number: A37549; MUID:77193964; PMID:266717

A:Accession: A37549

A:Molecule type: protein

A:Residues: 44-118, 'N', 120, 'S', 122-163, 'I', 165-175, 'A', 177-182, 'T', 184-193, 'MV', 196-308,

R:Butkowski, R.J.; Ellison, J.; Downing, M.R.; Mann, K.G.

J. Biol. Chem. 252, 4942-4957, 1977

A:Title: Primary structure of human prothrombin 2 and alpha-thrombin.

A:Reference number: A37550; MUID:77207112; PMID:873923

A:Accession: A37550

A:Molecule type: protein

A:Residues: 315-334, 'N', 336-348, 'N', 350-368, 'N', 370-397, 'N', 399-413, 'N', 415-484, 'N', 486-

R:Rabiet, M.J.; Blaschill, A.; Furie, B.; Furie, B.C.

Biochemistry 26, 13210-13215, 1986

A:Reference number: A37551; MUID:87008532; PMID:3759558

Contents: annotation; activation cleavages

MacGillivray, R.T.; Irwin, D.M.; Guinto, E.R.; Stone, J.C.

Ann. N. Y. Acad. Sci. 485, 73-79, 1986

A:Title: Recombinant genetic approaches to functional mapping of thrombin.

A:Reference number: I51952; MUID:87182874; PMID:34711151

A:Accession: I51952

A:Molecule type: mRNA

A:Residues: 1-2, 'RI', 5-100 &lt;RES&gt;

A:CROSS-references: GB:M33031; NID:g190723; PIDN:AAA60220.1; PID:g190724

C:Comment: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fibrin

) can be removed either by factor Xa or thrombin; the cleavage into light and heavy chain

ter 314-Arg, are released in natural blood clotting

C:Comment: The cleavage after Arg-198, observed in vitro, does not occur in plasma.

C:Comment: The gamma-carboxyglutamyl residues bind calcium ions, result from the carboxy

ent interaction with the negatively charged phospholipid membrane surface.

C:Comment: The prothrombin precursor is synthesized in the liver.

C:Genetics:

A:Gene: GDB:F2

A:CROSS-references: GDB:I19894; OMIM:176930

A:Map position: 11p11-11q12

A:Introns: 27/1; 80/3; 89/1; 106/1; 141/2; 187/1; 292/1; 335/1; 377/2; 433/2; 491/2; 552

C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology

C:Keywords: acute phase; blood coagulation; calcium binding; carboxyglutamic acid; dupli

F;1-24/Domain: signal sequence #status predicted &lt;SIG&gt;

F;25-43/Domain: propeptide #status predicted <PRO>  
F;28-87/Domain: Gla domain homology <GLA>  
F;44-622/Product: prothrombin #status experimental <MAT>  
F;44-327/Domain: activation peptide #status experimental <APT>  
F;108-186/Domain: kringle homology <KRI>  
F;213-291/Domain: kringle homology <KR2>  
F;328-363/Product: thrombin light chain #status experimental <LCH>  
F;364-622/Product: thrombin heavy chain #status experimental <HC>  
F;364-613/Domain: trypsin homology <TRY>  
F;49,50,57,59,62,63,68,69,72,75/Modified site: gamma-carboxyglutamic acid (Glu) #sta

F;60-65,90-103,108-186,129-169,157-181,213-291,234-274,262-286/Disulfide bonds: #sta

F;121,143/Binding site: carboxyhydrate (Asn) (covalent) #status predicted

F;336-407/Disulfide bonds: #status predicted

F;391-407/Disulfide bonds: #status experimental

F;406,462/Active site: His, Asp #status predicted

F;416/Binding site: carboxyhydrate (Asn) (covalent) #status experimental

F;568/Active site: Ser #status experimental

Alignment Scores:

Pred. No.: 1.79e-27 Length: 622

Score: 562.50 Matches: 169

Percent Similarity: 39.76% Conservative: 62

Best Local Similarity: 29.09% Mismatches: 161

Query Match: 24.07% Indels: 189

DB: 1 Gaps: 24

US-09-763-153-1 (1-1245) x TBHU (1-622)

QY 1 GCACACTCTCTCTCGAGAGCTCGTCACACAGCCCTGGAGCGGAGTGATAGAGAG 60

Db 44 AlaAsnThrPheLeuGluValArgLysGlyAsnLeuGluArgGluGluValGluGlu 63

QY 61 ATCTGTGACTTGGAGAGCGCCAAAGAAATTTCCAAATATGGATGACACACTGGCCCTC 120

Db 64 ThrCysSerTyrGluAlaPheGluAlaLeuGluSerThrAlaThrAspValPhe 83

QY 121 TGGTCCACAGCAGTCGACAGTGTCAGTGCTTGGCTTGGCTTGGAGCAGCCGTCGGCC 180

Db 84 TrpAlaLysTyrThrAlaCysGluThrAlaArgThr---ProArgAspLysLeuAlaAla 102

QY 181 AGCCTGTCTCGGCGCAGCAGCTGCATGCAGCGCATCGCAGCAGCAGCAGCAGCAGC 225

Db 103 ---CysLeuGluGlyAsnCysAlaGluGlyLeuGlyThrAsnTyrArgGlyHis 119

QY 226 ---TTCAGCTGCCAC--- 237

Db 120 ValAsnIleThrArgSerGlyIleGluCysGlnLeuTrpArgSerArgTyrProHisLys 139

QY 238 ---TTCAGCTGCCAC---TGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 243

Db 140 ProGluIleAsnSerThrThrHisProGlyAlaAspLeuGlnGluAsnPheCysArgAsn 159

QY 244 ---AGCGGCTGGAGGCGCGCTTCTGC--- 267

Db 160 ProAspSerThrThrGlyProTyrCysTyrThrThrAspProThrValArgArgGln 179

QY 268 ---CAGCGC--- 273

Db 180 GluCysSerIleProValCysGlyGlnAspGlnValThrValAlaMetThrProArgSer 199

QY 274 GAGGTGAGCTTCCTCAAT---TGCTCGCTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 312

Db 200 GluGlySerValAsnLeuSerProProLeuGluGlnCysValProAspArgGlyGln 219

QY 313 ---TGCAGC 321

Db 220 GlnTyrGlnGlyArgLeuAlaValThrThrHisGlyLeuProCysLeuAlaTrpAlaSer 239

QY 321 --- 321

Db 240 AlaGlnAlaLysAlaLeuSerLysHisGlnAspPheAsnSerAlaValGlnLeuValGlu 259

QY 322 ---TACTGCCTA-----GAGAGGTTGGCTGGCGCGCTGTAGCTGTCCG 363







QY	10	TTCTCTGGAGGACTCCGTCACAGCAGCCTGGAGCGGAGTGCATAGAGAGATCTGTGAC	69
Db	48	PheLeuGluGluValArgGlyGlyAsnLeuGluArgGluCysLeuGluGluProCysSer	67
QY	70	TTCCAGAGGCCCAAGGAAATTTCCAAATGTGGATGACACACTGGCCTTCTGGTCCAA	129
Db	68	ArgGluGluAlaPheGluAlaLeuGluLuserLeuSerAlaThrAspAlaPheTrpAlaLys	87
QY	130	CAGCTCGAGCGGTGACCAAGTCTTGCTTGCCTTGGAGCACCCGTGCGCCAGCCTGTC	189
Db	88	TyrThrAlaCysGluSerAla-----ArgAsnProArgGluLysLeuAsn	102
QY	190	---TCGGCGCAGCGCAGTGCATCGACGGCATC-----GGCAGCTTCAGC	231
Db	103	GluCysLeuGluGlyAsnCysAlaGluGlyValGlyMetAsnTyrArgGlyAsnValSer	122
QY	232	TGCGACTCGCCAGCGGC-----TGGAGGCGCGC-----	261
Db	123	Val---ThrArgSerGlyIleGluCysGlnLeuTrpArgSerArgTyrProHisLysPro	141
QY	262	-----TTCTGCG-----	267
Db	142	GluIleAsnSerThrThrHisProGlyAlaAspLeuArgGluAsnPheCysArgAsnPro	161
QY	268	-----CAGCGCGAG	276
QY	162	AspGlySerIleThrGlyProTrpCysTyrThrThrSerProThrLeuArgGluGlu	181
Db	182	CysSerValProValCysGlyGlnAspArgValThrValGluValIleProArgSerGly	201
QY	310	GCTGTCAG-----	318
Db	202	GlySerThrThrSerGlnSerProLeuLeuGluThrCysValProAspArgGlyArgGlu	221
QY	318	-----	318
Db	222	TyrArgGlyArgLeuAlaValThrThrSerGlySerArgCysLeuAlaTrpSerSerGlu	241
QY	319	-----CAT	321
Db	242	GlnAlaLysAlaLeuSerLysAspGlnPheAsnProGluAlaValProLeuAlaGluAsn	261
QY	322	TACTGCGCTA-----GAGGAGTGGCGTGGCGCGCTGTAGCTGTGCGCGCT	366
Db	262	PheCysArgAsnProAspGlyAspGluGlyAlaTrpCysTyrValAlaAspGlnPro	281
QY	367	GGCTTCAAGCTGGGGAGCCACCTCTCGAGTGTACCCCCCAGTG-----	411
Db	282	GlyAspPheGluTyrCysAspLeuAsnTyrCysGluGluProValAspGlyAspLeuGly	301

C;Date: 14-Dec-1990 #sequence\_revision 14-Dec-1990 #text\_change 03-May-2002  
 C;Accession: A35827; A42696; S12081  
 R;Degen, S.J.F.; Schaefer, L.A.; Jamison, C.S.; Grant, S.G.; Fitzgibbon, J.J.; Pai, J.A.  
 DNA Cell Biol. 9, 487-498, 1990  
 A;Title: Characterization of the cDNA coding for mouse prothrombin and localization of  
 A;Reference number: A35827; MUID:91025551; PMID:2222810  
 A;Accession: A35827  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-618 <DEG>  
 A;Cross-references: GB:X52308; NID:g53813; PIDN:CAA36548.1; PID:g53814  
 A;Experimental source: strain C57BL/6  
 A;Note: the data were obtained from females resulting from the cross of M. domesticus an  
 R;Banfield, D.K.; MacGillivray, R.T.A.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992  
 A;Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and seq  
 A;Reference number: A42696; MUID:92212913; PMID:1557383  
 A;Accession: A42696  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 384-618, E' <BAN>  
 A;Cross-references: GB:M81394  
 A;Experimental source: thrombin; Gla domain homology; kringle homology; trypsin homology  
 A;Note: superfamily: thrombin; Gla domain homology; kringle homology <KR2>  
 A;Keywords: blood coagulation; calcium binding; carboxyglutamic acid; glycoprotein; hyd  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-43/Domain: propeptide #status predicted <PRO>  
 F:28-88/Domain: Gla domain homology <GLA>  
 F:44-618/Product: prothrombin B #status predicted <MAT>  
 F:109-187/Domain: kringle homology <KR1>  
 F:215-293/Domain: kringle homology <KR2>  
 F:361-610/Domain: trypsin homology <TRY>  
 F:50-51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #status  
 F:61-66,91-104,109-187,130-170,158-182,215-293,236-276,264-288,333-479,388-404,533-547,5  
 F:403,459,565/Active site: His, Asp, Ser #status predicted

Alignment Scores:  
 Pred. No.: 1,22e-25 Length: 618  
 Score: 533.00 Matches: 167  
 Percent Similarity: 39.59% Conservative: 65  
 Best Local Similarity: 28.50% Mismatches: 152  
 Query Match: 22.81% Indels: 207  
 DB: 2 Gaps: 2

US-09-763-153-1 (1-1245) x A35827 (1-618)

181 AGCCTGTGCTGGCGGCAC----- 198  
 184 ValProValCysGlyGlnGluGlyArgThrValValMetThrProArgSerGlyGly 203  
 199 -----GCACGTGCATCGACGGCATCGC----- 222  
 204 SerLysAspAsnLeuSerProProLeuGlyGlnCysLeuThrGluArgGlyArgLeuTyr 223  
 222 ----- 222  
 224 GlnGlyAsnLeuAlaValThrThrLeuGlySerProCysLeuProTrpAsnSerLeuPro 243  
 223 -----ACCTTCAGCTGGCAG----- 237  
 244 AlaLysThrLeuSerLysTyrGlnAspPheAspProGluValLysLeuValGluAsnPhe 263  
 238 TGCGCGCAGC---GGCTGG-----GAGGGCGGCTTCTGCGCAG-----CGCAGAGTG 279  
 264 CysArgAsnProAspTrpAspGluGlyAlaTrpCysTyrValAlaGlyGlnProGly 283  
 280 AGCTTCTCCTCAATTGCTGCTGGACAAACGGGGCTGCACCGCATTTACTGCTAGAGGAGTG 339  
 284 AspPheGluTyrCysAsnLeu-----AsnTyrCysGluGluAlaVal 297  
 340 GGCTGGCGCGCTGTAGCTGTGCGCTGCTACAAGCTGGGGCAGCAGCTCTCTGCAGTGT 399  
 298 Gly-----GluGluAsnTyrAspValAspGluSerIleAlaGlyArg 311  
 400 CACCCCGCAGTGAAGTTC----- 417  
 312 ThrThrAspAlaGluPheHisThrPhePheAsnGluLysThrPheGlyLeuGlyGluAla 331  
 418 CTTTGTGGG---AGGCCCTGGAGCGGATGGAGAGAGCGCAGTCACCTGAAACAGAGAC 474  
 332 AspCysGlyLeuArgPro-----LeuPheGluLys-----SerLeuLysAspThr 347  
 475 ACAGAACAGCAA-----GAAGACCAAGTAGATCCGCGCTCATGTGATGGAGAGATGACC 528  
 348 ThrGluLysGluLeuLeuLeuAspSerTyrIleAspGlyArgIleValGluGlyTrpAspAla 367  
 529 AGCGGGGAGACAGCCCTGCGAGGTGCTGCTGCTG---GACTCAAGAGAGAGCTGCC 585  
 368 GluLysGlyIleAlaProTrpGlnValMetLeuPheArgLysSerProGlnGluLeu 387  
 586 TCGGGGGCAGTGCATCCACCCCTCTGCTGGTGTGACAGCGGCCACTGTCATG----- 639  
 388 CysGlyAlaSerLeuIleSerAspArgTrpValLeuThrAlaAlaHisCysIleLeuTyr 407  
 640 -----GATGAGTCC-----AAGAAAGCTCTCTGTGTCAGCTTGGAGATATGAC 681  
 408 ProProTrpAspLysAsnPheThrGluAsnAspLeuLeuValArgIleGlyLysHisSer 427  
 682 CTGCGGCGCTGGGAGAAAG---TGGGAGCTGGACCTGGACATCAAGGAGTCTTCTCCAC 738  
 428 ArgThrArgTyrGluArgAsnValLysIleSerMetLeuGluLysIleTyrValHis 447  
 739 CCCAACTACAGC---AAGAGCACCCAGCAATGATCATCGCAGCTGCTCAGCTGCCAC 795  
 448 ProArgTyrAsnTrpArgGluAsnLeuAspArgAspIleAlaLeuLeuLysLys 467  
 796 CCGCGCCCTCTCCAGACCATAGTGCCTCTCTCCCGGACAGCGGCGCTTGCAGAG 855  
 468 ProValProPheSerAspTyrIleHisProValCysLeuProAspLysGlnThrVal--- 486  
 856 CCGAGACTCAATCAGCGCGCGCAGGAGACCCCTGCTGAGCGGCTGGGTCTACACAGCAGC 915  
 487 ThrSerLeuAlaGlyTyrArgValGlyArgValGlyTyrGly-----AsnLeu 504  
 916 CGAGAG-----AAGGAGGCCAAGAAACCCGACCTTCTGCTCACTTCAATCAAG 966  
 505 ArgGluThrTrpThrThrAsnIleAsnGluIleGlnProSerValLeuGlnValAsn 524  
 967 ATTCCCGTGGTCCCGCACCAATGAGTGCAGCGAGGTATCATGAGCAACATGGTCTCAGAAC 1026

Db 525 LeuProIleValGluArgProValCysLysAlaSerThrArgIleThrAspAsn 544  
 QY 1027 ATGCTGTGTGGGGGATC---CTCGGGGAC-----CGGAGGATGCTCGGAGGGGAC 1077  
 Db 545 MetPheCysAlaGlyPheLysValAsnAspThrLysArgGlyAspAlaCysGluGlyAsp 564  
 QY 1078 ATGGGGGGCCCATGGTC-----GCCCTCTCCACGGCACCTGGTTCCTGGTGGGCCCTG 1131  
 Db 565 SerGlyGlyProPheValMetLysSerProPheAsnAsnArgTrpTyrGlnMetGlyIle 584  
 QY 1132 GTGAGCTGGGTGAGGGCTGTGGCTCCTTCACAACTACGGCTTTACACCAAGTCACG 1191  
 Db 585 ValSerTrpGlyGluGlyCysAspArgLysGlyLysTyrGlyPheTyrThrHisValPhe 604  
 1192 CGTACCTACCTGCTGATC 1209  
 605 ArgLeuLysArgTrpIle 610

## RESULT 19

S10511  
 thrombin (EC 3.4.21.5) precursor - rat  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 03-May-2002  
 R;Accession: S10511; A60576; B42696  
 R;Dihaneich, M.; Monard, D.  
 Nucleic Acids Res. 18, 4251, 1990  
 A;Title: cDNA sequence of rat prothrombin.  
 A;Reference number: S10511; MUID:90332426; PMID:2377469  
 A;Accession: S10511  
 A;Molecule type: mRNA  
 A;Residues: 1-617 <DIR>  
 A;Cross-references: EMBL:X52835; NID:956969; PID:CAA37017.1; PID:956970  
 R;Henrikson, K.P.; Jazin, E.E.; Greenwood, J.A.; Dickerman, H.W.  
 Endocrinology 126, 167-175, 1990  
 A;Title: Prothrombin levels are increased in the estrogen-treated immature rat uterus.  
 A;Reference number: A60576; MUID:90091942; PMID:2293980  
 A;Accession: A60576  
 A;Molecule type: protein  
 A;Residues: 44-58 <HEN>  
 A;Note: the authors purified the proenzyme from the estrogen-stimulated maturing rat ute  
 R;Banfield, D.K.; MacGillivray, R.T.A.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992  
 A;Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and seq  
 A;Reference number: A42696; MUID:92212913; PMID:1557383  
 A;Accession: B42696  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 383-617, 'E' <BAN>  
 A;Cross-references: GB:M81397  
 C;Superfamily: thrombin; Gla domain homology; kringe homology; trypsin homology  
 C;Keywords: blood coagulation; calcium binding; kringe homology; trypsin homology  
 F;1-24/Domain: signal sequence #status predicted <SIG>  
 F;25-43/Domain: propeptide #status predicted <PRO>  
 F;44-617/Product: prothrombin #status experimental <PMAT>  
 F;109-187/Domain: kringe homology <KR1>  
 F;215-292/Domain: kringe homology <KR2>  
 F;360-609/Domain: trypsin homology <TRY>  
 F;50-51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxyglutamic acid (Glu) #status  
 F;61-66,91-104,109-187,130-170,158-182,215-292,326-276,264-287,332-478,387-403,532-546,5  
 F;402,458,564/Active site: His, Asp, Ser #status predicted

## Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
3.57e-25	525.50	617	165
Percent Similarity:	40.21%	Conservative:	65
Best Local Similarity:	28.85%	Mismatches:	167
Query Match:	22.45%	Indels:	175
DB:	2	Gaps:	24

US-09-763-153-1 (1-1245) x S10511 (1-617)

QY	1	GCCAACTCC	---	TTCTGGAGGAGCTCGTCCACAGCAGCTGGAGCGGAGTGCATAGAC	57
Db	44	AlaAsnSerGlyPheLeuGluGluLeuArgLysGlyAsnLeuGluArgGluCysValGlu	63		
QY	58	GAGATCTCTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGGATGACACACTGGCC	117		
Db	64	GluGlnCysSerTyrGluGluAlaPheGluAlaLeuGluSerProGlnAspThrAspVal	83		
QY	118	TTCTGGTCCAAAGCAGCTCGACGGTGAC	---	---	CAG 147
Db	84	PheTrpAlaLysTyrThrValCysAspSerValArgLysProArgGluThrPheMetAsp	103		
QY	148	TGCTTG	---	---	153
Db	104	CysLeuGluGlyArgCysAlaMetAspLeuGlyLeuAsnTyrHisGlyAsnValSerVal	123		
QY	154	---	---	---	GTC 156
Db	124	ThrHisThrGlyIleGluCysGlnLeuTrpArgSerArgTyrProHisArgProAspIle	143		
QY	157	TTGCCCTTGAGCAGCCCGTCCGCCAGCCCTG	---	---	TGCTGGGGGACGGCAGC 204
Db	144	AsnSerThrThrHisProGlyAlaAspLeuLysGluAsnPheCysArgAsnProAspSer	163		
QY	205	TGCATCAGCGGATCGGCAGCTTACAGTCGCAC	---	---	TGCCGC 243
Db	164	SerThrSerGlyProTrpCysTyrThrAspProThrValArgGluGluCysSer	183		
QY	244	---	---	---	---
Db	184	IleProValCysGlyGlnGluGlyArgThrThrValLysMetThrProArgSerArgGly	203		
QY	295	TCGCTGGCAAC	---	---	315
Db	204	SerLysGluAsnLeuSerProLeuGlyGluCysLeuLeuGluArgGlyArgLeuTyr	223		
QY	315	---	---	---	---
Db	224	GlnGlyAsnLeuAlaValThrThrLeuGlySerProCysLeuAlaTrpAspSerLeuPro	243		
QY	316	---	---	---	AGCATTAC 324
Db	244	ThrLysThrLeuSerLysTyrGlnAsnPheAspProGluValLysLeuValGlnAsnPhe	263		
QY	325	TGCCTA	---	---	---
Db	264	CysArgAsnProAspArgAspGluGluGlyAlaTyrCysPheValAlaGlnGlnProGly	283		
QY	370	TACAAGCTGGGGACGACCTCTGCAGTCACCCCGCAGTC	---	---	411
Db	284	PheGluTyr	---	---	---
QY	411	---	---	---	---
Db	303	GlyAspGluSerIleAlaGlyArgThrThrAspAlaGluPheHisThrPhePheAspGlu	322		
QY	412	---	---	---	AGCCCTCGAAGCGGAGGAG 450
Db	323	ArgThrPheGlyLeuGlyGluAlaAspCysGlyLeuArgPro	340		
QY	451	AACCGGAGTCACCTGAACAGACACAGACACCAAGAACCAAGATAGATCCCGGCTC	510		
Db	341	LysSerLeuThrAspLysThrGluLysGluLeuLeuAspSerTyrIleAspGlyArgIle	360		
QY	511	ATTGATGGGAAGATACACAGCGGGGACACAGCCCTCGAGGTGGTCTCTCTG	567		
Db	361	ValGluGlyTrpAspAlaGluLysGlyIleAlaProTrpGlnValMetLeuPheArgLys	380		
QY	568	TCAAGAAAGAGCTGGCTGGCGGCGAGTCTCATCCACCCCTCTGGGTGCTGACAGCG	627		
Db	381	SerProGlnGluLeuLeuCysGlyAlaSerLeuLeuSerAspArgTrpValLeuThra	400		
QY	628	GCCCACTGCATG	---	---	AAGAAGCTCTTGTC 663

A:Accession: F01268  
 A:Author: Miyazawa, K.; Shimomura, T.; Iwata, J.; Biol. Chem. 268, 10024-10028, 1993  
 A:Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease  
 A:Abstract: The cDNA for a human serine protease was cloned and sequenced. The deduced  
 A:Keywords: Molecular cloning; sequence analysis; serine protease  
 A:Publication factor XII.  
 A:Reference number: A46688; MUID:93252878; PMID:7683665  
 A:Accession: A46688  
 A:Molecule type: mRNA  
 A:Residues: 1-655 <M>Y  
 A:Cross-references: DDBJ:D14012; NID:g219680; PIDN:BAA03113.1; PID:g219681  
 A:Experimental source: liver (mRNA); serum (protein)  
 A:Note: sequence extracted from NCBI backbone (NCBIN:131227, NCBIP:131228)  
 A:Note: parts of the sequence, including the amino ends of the heavy and light chains, c  
 C:Genetics:  
 A:Gene: GDB:HGFAC; HGFA; HGFAP  
 A:Cross-references: GDB:9954514  
 A:Map position: 4p16-4p16  
 C:Function:  
 A:Description: activates hepatocyte growth factor by specific proteolytic cleavage  
 A:Pathway: tissue repair and regeneration  
 C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology  
 C:Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase  
 F:1-34/Domain: signal sequence #status predicted <SIC>  
 F:108-148/Domain: fibronectin type II repeat homology <1F2>  
 F:164-197/Domain: EGF homology <EG1>

QY	43	CGGAGTGTAGAGAGATCTGTGACTTCGAGGAGGCCAAGGAATATTTCCAAATGTG	102
Db	195	LysAspCysGlyThrGluLysCysPheAsnGluThrArgTyrGluTyrLeuGluGlyGly	214
QY	103	GATGCACACTGGCCTTCTGGTCCCAAGCAGCTCGACGGTGACACAGCTGTG	153
Db	215	AspArgTrpAlaArgValArgLncGlyHisValGluLncCysLysCysPheGlyGlyArg	234
QY	154	-----GTCTTGGCCTTGGAGCACCCCTGGCTCCGACCGCTGTGC	189
Db	235	ThrTrpCysGluGlyThrArgHisThrAlaCysLeuSerSerProCysLeuAsn	252
QY	190	TGCGGCAGCGCAGTGC-----ATCCAGCGCATCGGCAGCTTCAGCTCGCAGTCCGCC	243
Db	253	-----GlyGlyThrCysHisLeuIleValAlaThrGlyThrValCysAlaCysPro	270
QY	244	AGCGGCTGGGAGGCGCGCTCTTCGACGCGGAGGTGAGCTTCCTCAATTGCTCGCTGGAC	303
Db	271	ProGlyPheAlaGlyArgLeuCysAsnIleGluProAsp---GluArgCysPheLeuGly	289
QY	304	AACGGC-----	309
Db	290	AsnGlyThrGlyTyrArgGlyValAlaSerThrSerAlaSerGlyLeuSerCysLeuAla	309
QY	309	-----	309
Db	310	TrpAsnSerAspLeuLeuTyrGlnGluLeuHisValAspSerValGlyAlaAlaLeu	329
QY	310	-----GCCTCACGCATTACTGCCTA-----GAGGAGTGGGTGG---	345
Db	330	LeuGlyLeuGlyProHisAlaTyrCysArgAsnProAspAsnAspGluArgProTrpCys	349
QY	346	-----CGCGCTGTAGTGT	360
Db	350	TyrValValLysAspSerAlaLeuSerTrpGluTyrCysArgLeuGluAlaCysGluSer	369
QY	361	GGCCCTGGCTACAGCTGGGGAGACCTCCTCCAGTGTACCCCGCGAGTGAAGTTCCT	420
Db	370	LeuThrArgValGlnLeuSerProAspLeuLeuAlaThrLeuProGluProAlaSerPro	389
QY	421	TGTGGGAGGCCCTTGAAGCGGATGGAAGAGCGGAGTCACCTGAACGAGACACAGAA	480
Db	390	GlyArgGlnAlaCysGlyArgArgHisLysLysArgThrPheLeuArg	405
QY	481	GACCAGAGACCAAGTAGATCCGCGGCTCATTCGGGAAGATGACACGCGGGGAGAC	540
Db	406	-----ProArgIleIleIleGlySerSerLeuProGlySer	418
QY	541	AGCCCTGG-----CAGGTGTCTGTCTGACTCAAAAGAAGCTGCGCTCGCGGGCA	594
Db	419	HisProTrpLeuAlaAlaIleIleGlyAspSer-----PheCysAlaGly	434
QY	595	GTGCTCATCACCCCTTCCTGGGTGCTGACGCGGCCCATGCTGATGATGATGCC-----	648

Db	435	SerLeuValHisThrCysTrpValValSerAlaAlaHisCysPheSerHisSerPro	454
QY	649	---AAGAAGCTCTTGTGAGGTATGACCTGGCGCTGGGAGAGTGGGAG	705
Db	455	ArgAspSerValSerValLeuGlyGlnHisPheAsnArgThrAspValThr	474
QY	706	CTGGACCTGGACATCAGGAGTCTCGTCCACCCCACTACAGCAAG---	762
Db	475	GlnThrPheGlyIleGluLeuTyrIleProTyrThrLeuTyrSerValPheAsnProSer	494
QY	763	GACATGACATCGACCTGCTGCACCTGCCCCAGCC-----CCGACCCTCTCG	810
Db	495	AspHisAspLeuValLeuIleArgLeuLysLysGlyAspArgCysAlaThrArgSer	514
QY	811	CAGACCATAGTGGCCATCTGCCCTCCGGACAGCGGGCTTGCAGAGCGGAGCTCAATCAG	870
Db	515	GlnPheValGlnProIleCysLeuProGluProGly-----SerThrPhePro	530
QY	871	CCCGCGCAGGAGACCTCGTGCAGCGGCTGGGGCTAC-----CAC	909
Db	531	AlaGlyHisLysCysGlnIleAlaGlyTrpGlyHisLeuAspGluAsnValSerGlyTyr	550
QY	910	AGCAGCGCAGAGAGGAGCGCCAGAACCCGACCTTCGTCTCAACTCATCAAGATT	969
Db	551	SerSerSerLeuArgGluAla-----LeuVal	559
QY	970	CCGTGTGTCGCCAATGAGTGCAC-----GAGGTCATGACCACTGGTCTCTGAG	1023
Db	560	ProLeuValAlaAspHisLysCysSerSerProGluValTyrGlyAlaAspIleSerPro	579
QY	1024	AACATGCTGTGCGGGCATCTCCGGGACCGCAGGATGCTCGAGGCGCAGAGTGGG	1083
Db	580	AsnMetLeuCysAlaGlyTyrPheAspCysLysSerAspAlaCysGlnGlyAspSerGly	599
QY	1084	GGCCCATGGTCCGCTCTTCACGGCACCTCGTGGTCTGGTGGCGCTGGTGGCTGGGT	1143
Db	600	GlyProLeuAlaCysGluLysAsnGlyValAlaTyrLeuTyrGlyIleIleSerTrpGly	619
QY	1144	GAGGCTGTGGGCTCTCTCACAACTACGGCTTTACACCAAGTCAGCCGCTACCTCGAC	1203
Db	620	AspGlyCysGlyArgLeuHisLysProGlyValTyrThrArgValAlaAsnTyrValasp	639
QY	1204	TGGATCCATGGCCACATCAGA	1224
Db	640	TrpIleAsnAspArgIleArg	646

Search completed: March 17, 2003, 17:37:13  
Job time : 70 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: March 17, 2003, 17:01:24 ; Search time 40 Seconds  
(without alignments)  
2581.901 Million cell updates/sec

Title: US-09-763-153-1

Perfect score: 2337

Sequence: 1 gccaaactctctctggagga.....acaaggaagccccccagaag 1245

ring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 60 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-O=cg2\_1/USPTO.spool/US09763153/runat\_11032003\_084247\_1981/app-query.fasta\_1.1415  
-DB=SwissProt\_40 -QMT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOPCL=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi  
-LIST=60 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=20  
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09763153 -CGN 1.1.16 -runat\_11032003\_084247\_1981 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG-SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2298	98.3	461	1 PRTC_HUMAN	P04070 homo sapien
2	1872.5	80.1	458	1 PRTC_RABIT	Q28661 oryctolagus
3	1750.5	74.9	459	1 PRTC_PIG	Q981p2 sus scrofa
4	1660	71.0	456	1 PRTC_BOVIN	P00745 bos taurus
5	1650.5	70.6	461	1 PRTC_RAT	P31394 rattus norv
6	1637.5	70.1	461	1 PRTC_MOUSE	P33587 mus musculus
7	835	35.7	161	1 PRTC_MACMU	Q28506 macaca mula
8	814.5	34.9	490	1 PRTC_HUMAN	O19045 oryctolagus
9	809	34.6	492	1 FA10_BOVIN	P00743 bos taurus
10	802	34.3	444	1 FA7_RABIT	P98139 oryctolagus
11	801.5	34.3	475	1 FA10_CHICK	P25155 gallus gall
12	801.5	34.3	488	1 FA10_HUMAN	P00742 homo sapien
13	783	33.5	466	1 FA7_HUMAN	P08709 homo sapien
14	779.5	33.4	407	1 FA7_BOVIN	P22457 bos taurus
15	770	32.9	446	1 FA7_MOUSE	P70375 mus musculus
16	763	32.6	452	1 FA9_CANFA	P19540 canis famil
17	736	31.5	461	1 FA9_HUMAN	P00740 homo sapien
18	726	31.1	459	1 FA9_MOUSE	P16294 mus musculus

19	724	31.0	376	1	FA10_TROCA	P81428 tropidechis
20	717	30.7	157	1	PRTC_CANFA	Q28278 canis famil
21	716	30.6	157	1	PRTC_FELCA	Q28412 felis silve
22	714.5	30.6	416	1	FA9_BOVIN	P00741 bos taurus
23	700	30.0	157	1	PRTC_HORSE	Q28380 equus cabal
24	661	28.3	157	1	PRTC_CAPHI	Q28315 capra hircu
25	562.5	24.1	622	1	THRB_HUMAN	P00734 homo sapien
26	538.5	23.0	625	1	THRB_BOVIN	P00735 bos taurus
27	533	22.8	618	1	THRB_MOUSE	P19221 mus musculus
28	525.5	22.5	617	1	THRB_RAT	P18292 rattus norv
29	475.5	20.3	653	1	HGFA_MOUSE	Q09098 mus musculus
30	470	20.1	655	1	HGFA_HUMAN	Q04756 homo sapien
31	468	20.0	400	1	PRTC_HUMAN	P22891 homo sapien
32	445.5	19.1	396	1	PRZ_BOVIN	P00744 bos taurus
33	439.5	18.8	275	1	TRVT_PIG	Q9nzd1 sus scrofa
34	427	18.3	271	1	FA9_PIG	P16293 sus scrofa
35	426	18.2	638	1	KAL_MOUSE	P26262 mus musculus
36	424.5	18.2	638	1	KAL_HUMAN	P03952 homo sapien
37	421	18.0	455	1	TMS5_MOUSE	Q9er04 h complemen
38	420.5	18.0	699	1	CRAR_HUMAN	P16292 oryctolagus
39	420	18.0	275	1	FA9_RABIT	P20918 mus musculus
40	419.5	18.0	812	1	PLMN_MOUSE	P03951 homo sapien
41	419	17.9	625	1	FA11_HUMAN	Q9y1q8 mus musculus
42	418	17.9	490	1	TMS2_MOUSE	P56677 mus musculus
43	417.5	17.9	855	1	ST14_MOUSE	P16295 cavia porce
44	417	17.8	285	1	FA9_CAVPO	P98064 mus musculus
45	415.5	17.8	704	1	CRAR_MOUSE	P14272 rattus norv
46	414	17.7	638	1	KAL_RAT	Q05319 drosophila
47	411	17.6	786	1	STUB_DROME	P97435 mus musculus
48	411	17.6	1069	1	ENTK_MOUSE	P00747 homo sapien
49	410	17.5	810	1	PLMN_HUMAN	P16296 rattus norv
50	409.5	17.5	282	1	FA9_RAT	P56730 homo sapien
51	408.5	17.5	875	1	NETR_HUMAN	Q9p0g3 homo sapien
52	406	17.4	251	1	KLKE_HUMAN	P05981 homo sapien
53	406	17.4	417	1	HEPS_HUMAN	O15393 homo sapien
54	405.5	17.4	492	1	TMS2_HUMAN	O15661 homo sapien
55	404.5	17.3	275	1	TRB1_HUMAN	P28175 tachypleus
56	404	17.3	1019	1	LFC_TACTR	P20231 homo sapien
57	403.5	17.3	275	1	TRB2_HUMAN	Q9xsm2 ovis aries
58	403	17.2	273	1	TRVT_SHEEP	P05156 homo sapien
59	403	17.2	583	1	CEAL_HUMAN	P15157 homo sapien
60	402.5	17.2	275	1	TRYA_HUMAN	

#### ALIGNMENTS

RESULT 1  
PRTC\_HUMAN  
ID PRTC\_HUMAN STANDARD; PRT: 461 AA.  
AC P04070; Q16001; Q15190; Q15189;  
DT 01-NOV-1986 (Rel. 03, Created)  
DT 01-NOV-1986 (Rel. 03, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Vitamin-K dependent protein C precursor (EC 3.4.21.69)  
DE (Autoproteolysin IIA) (Anticoagulant protein C) (Blood coagulation factor XIV).  
DE PROC.  
GN Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85270390; PubMed=2991887;  
RA Foster D.C., Yoshitake S., Davie E.W.;  
RT "The nucleotide sequence of the gene for human protein C.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:4673-4677(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85296939; PubMed=2991859;  
RA Beckmann R.J., Schmidt R.J., Santerre R.F., Plutzky J., Crabtree G.R.,  
RA Long G.L.;  
RT "The structure and evolution of a 461 amino acid human protein C

precursor and its messenger RNA, based upon the DNA sequence of  
cloned human liver cDNAs.";  
Nucleic Acids Res. 13:5233-5247(1985).  
[3]  
SEQUENCE FROM N.A.  
MEDLINE-91329836; PubMed-1868249;  
RA Reitsma P.H., Poort S.R., Allaart C.F., Briet E., Bertina R.M.;  
RA Plutzky J., Hoskins J.A., Long G.L., Crabtree G.R.;  
RA "Evolution and organization of the human protein C gene.";  
Proc. Natl. Acad. Sci. U.S.A. 83:546-550(1986).  
[4]  
SEQUENCE FROM N.A.  
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,  
RA Nickerson D.A.;  
RA Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
[5]  
SEQUENCE OF 106-461 FROM N.A.  
RX MEDLINE-84272714; PubMed-6589623;  
RX Foster D.C., Davie E.W.;  
RA "Characterization of a cDNA coding for human protein C.";  
Proc. Natl. Acad. Sci. U.S.A. 81:4766-4770(1984).  
[6]  
CARBOHYDRATE-LINKAGE SITE ASN-371.  
MEDLINE-90293094; PubMed-1694179;  
RA Miletich J.P., Broze G.J. Jr.;  
RA "Beta protein C is not glycosylated at asparagine 329. The rate of  
translation may influence the frequency of usage at asparagine-X-  
cysteine sites.";  
J. Biol. Chem. 265:11397-11404(1990).  
[7]  
HYDROXYLATION.  
RX MEDLINE-92184750; PubMed-1544894;  
RA Harris R.J., Ling V.T., Spellman M.W.;  
RA "O-linked fucose is present in the first epidermal growth factor  
domain of factor XII but not protein C.";  
J. Biol. Chem. 267:5102-5107(1992).  
[8]  
3D-STRUCTURE MODELING OF 175-450.  
RX MEDLINE-94272342; PubMed-8003577;  
RA Fisher C.L., Greengard J.S., Griffin J.H.;  
RA "Models of the serine protease domain of the human antithrombotic  
plasma factor activated protein C and its zymogen.";  
Protein Sci. 3:588-599(1994).  
[9]  
X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 84-461.  
RX MEDLINE-97157472; PubMed-9003757;  
RA Mather I., Oganessyan V., Hof P., Huber R., Foundling S., Esmon C.,  
RA Bone M.;  
RA "The 2.8 A crystal structure of Gla-domainless activated protein C.";  
EMBO J. 15:6822-6831(1996).  
[10]  
REVIEW ON PROC VARIANTS.  
RX MEDLINE-93190290; PubMed-8446940;  
RA Reitsma P.H., Poort S.R., Bernardi F., Gandrille S., Long G.L.,  
RA Sala N., Cooper D.N.;  
RA "Protein C deficiency: a database of mutations. For the Protein C & S  
Subcommittee of the Scientific and Standardization Committee of the  
International Society on Thrombosis and Haemostasis.";  
Thromb. Haemost. 69:77-84(1993).  
[11]  
VARIANT CYS-444.  
RX MEDLINE-87204221; PubMed-2437584;  
RA Romeo G., Hassan H.J., Staepfli S., Roncuzzi L., Cianetti L.,  
RA Leonardi A., Vicente V., Mannucci P.M., Bertina R.M., Peschle C.,  
RA Cortese R.;  
RA "Hereditary thrombophilia: identification of nonsense and missense  
mutations in the protein C gene.";  
Proc. Natl. Acad. Sci. U.S.A. 84:2829-2832(1987).  
[12]  
VARIANT TRP-211 (LONDON-1).  
RX MEDLINE-90098906; PubMed-2602169;  
RA Grundy C.B., Chitologie A., Talbot S., Bevan D., Kakkar V.V.,  
RA Cooper D.N.;  
RA "Protein C London 1: recurrent mutation at Arg-169 (CGG-->TGG) in

the protein C gene causing thrombosis.";  
Nucleic Acids Res. 17:10513-10513(1989).  
[13]  
VARIANT CYS-272.  
RX MEDLINE-91329836; PubMed-1868249;  
RA Reitsma P.H., Poort S.R., Allaart C.F., Briet E., Bertina R.M.;  
RA "The spectrum of genetic defects in a panel of 40 Dutch families with  
symptomatic protein C deficiency type I: heterogeneity and founder  
effects.";  
Blood 78:890-894(1991).  
[14]  
VARIANTS ALA-62 (VERMONT-1) AND MET-76.  
RX MEDLINE-92190481; PubMed-1347706;  
RA Davila E.G., Tomczak J.A., Grant B., Bhushan F., Pillemer E.,  
RA Rainville I.R., Long G.L.;  
RA "Protein C Vermont: symptomatic type II protein C deficiency  
associated with two GLA domain mutations.";  
Blood 79:1456-1465(1992).  
[15]  
VARIANT ASP-418 (HONG KONG-2).  
RX MEDLINE-92305321; PubMed-1611081;  
RA Sugahara Y., Miura O., Yuen P., Aoki N.;  
RA "Protein C deficiency Hong Kong 1 and 2: hereditary protein C  
deficiency caused by two mutant alleles, a 5-nucleotide deletion and  
a missense mutation.";  
Blood 80:126-133(1992).  
[16]  
VARIANT LEU-289.  
RX MEDLINE-92380660; PubMed-1511988;  
RA Grundy C.B., Chisholm M., Kakkar V.V., Cooper D.N.;  
RA "A novel homozygous missense mutation in the protein C (PROC) gene  
causing recurrent venous thrombosis.";  
Hum. Genet. 89:683-684(1992).  
[17]  
VARIANTS GLN-220 AND TRP-220.  
RX MEDLINE-92380661; PubMed-1511989;  
RA Grundy C.B., Schulman S., Tengborn L., Kakkar V.V., Cooper D.N.;  
RA "Two different missense mutations at Arg 178 of the protein C (PROC)  
gene causing recurrent venous thrombosis.";  
Hum. Genet. 89:685-686(1992).  
[18]  
VARIANT GLN-220.  
RX MEDLINE-93250852; PubMed-1301959;  
RA Gandrille S., Vidaud M., Aiach M., Alhenc-Gelas M., Fischer A.M.,  
RA Couault-Hellman M., Toulon P., Flessinger J.N., Goossens M.;  
RA "Two novel mutations responsible for hereditary type I protein C  
deficiency: characterization by denaturing gradient gel  
electrophoresis.";  
Hum. Mutat. 1:491-500(1992).  
[19]  
VARIANT SER-334.  
RX MEDLINE-92276939; PubMed-1593215;  
RA Yamamoto K., Matsushita T., Suglura I., Takamatsu J., Iwasaki E.,  
RA Wada H., Deguchi K., Shirakawa S., Saio H.;  
RA "Homozygous protein C deficiency: identification of a novel missense  
mutation that causes impaired secretion of the mutant protein C.";  
J. Lab. Clin. Med. 119:682-689(1992).  
[20]  
VARIANTS TRP-38; CYS-42; HIS-42; GLN-271 AND ASN-294.  
RX MEDLINE-93313192; PubMed-8324221;  
RA Gandrille S., Alhenc-Gelas M., Gaussem P., Aillaud M.-F., Dupuy E.,  
RA Tjani-Vague I., Aiach M.;  
RA "Five novel mutations located in exons III and IX of the protein C  
gene in patients presenting with defective protein C anticoagulant  
activity.";  
Blood 82:159-168(1993).  
[21]  
VARIANTS G-14; Q-211; Y-244; Q-253; L-321; C-328; I-385; T-388 AND  
V-388.  
RX MEDLINE-93271391; PubMed-8499565;  
RA Poort S.R., Fabinger-Fasching I., Mannhalter C., Reitsma P.H.,  
RA Bertina R.M.;  
RA "Twelve novel and two recurrent mutations in 14 Austrian families



with hereditary protein C deficiency.";  
 RL Blood Coagul. Fibrinolysis 4:273-280(1993).  
 [22]  
 RP VARIANT TRP-57  
 RX MEDLINE=93271396; PubMed=8499568;  
 RA Miller D.S., Grundy C.B., Bignell P., Moffat E.H., Martin R.,  
 RA Kakkar V.V., Cooper D.N.;  
 RT "A Gla domain mutation (Arg 15-->Trp) in the protein C (PROC) gene  
 RT causing type 2 protein C deficiency and recurrent venous  
 RT thrombosis";  
 RL Blood Coagul. Fibrinolysis 4:345-347(1993).  
 [23]  
 RP VARIANTS R-145; L-210; W-211; T-243; L-321; M-340 AND Y-426.  
 MEDLINE=94122329; PubMed=8292730;  
 RA Tsay W., Greengard J.S., Montgomery R.R., McPherson R.A., Fucci J.C.,  
 RA Koepfer M.A., Coughlin J., Griffin J.H.;  
 RT "Genetic mutations in ten unrelated American patients with  
 RT symptomatic type 1 protein C deficiency.";  
 RL Blood Coagul. Fibrinolysis 4:791-796(1993).  
 [24]  
 RP VARIANT SER-423.  
 RX MEDLINE=94001606; PubMed=8398832;  
 RA Marchetti G., Patrachini P., Gemmati D., Castaman G., Rodeghiero F.,  
 RA Wacey A., Cooper D.N., Tuddenham E.G., Bernardi F.;  
 RT "Symptomatic type II protein C deficiency caused by a missense  
 RT mutation (Gly 381-->Ser) in the substrate-binding pocket.";  
 RL Br. J. Haematol. 84:285-289(1993).  
 [25]  
 RP SEQUENCE OF 43-64 FROM N.A., AND VARIANT GLY-57 (YONAGO).

## Alignment Scores:

Pred. No.: 1,41e-114 Length: 461  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 1 Gaps: 0

US-09-763-153-1 (1-1245) x PRTC\_HUMAN (1-461)

QY 1 GCCAACTCTTCCTGGAGAGCTCCGTACAGCAGCTGGAGCGGAGTGCATAGAGAG 60  
 DB 43 AlaAsnSerPheLeuGluGluLeuAargHisSerSerLeuGluAargGluCysileGluGlu 62  
 QY 61 ATCTGTGACTTCGAGAGGCCAAGAAATTTCCAAATATGTGATGACACACTGGCCTTC 120  
 DB 63 IleCysAspPheGluGluAlaLysGluilePheGlnAsnValaspAspThrLeuAlaPhe 82  
 QY 121 TGGTCCAGACGTCGACGCTGACCATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 180  
 DB 83 TrpSerLysHisValaspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
 QY 181 AGCCTGTGCTCGGGGACGCGACATGTCATCGACGCTGCGAGCTTCAGCTGCGACTGC 240  
 DB 103 SerLeuCysCysGlyHisGlyThrCysileaspGlyileGlySerPheSerCysaspCys 122  
 QY 241 CGCAGCGCTGGAGGCGCGCTTCTGCCAGCGGAGTGAGTTCCTCAATGCTCGCTG 300  
 DB 123 ArgSerGlyTrpGluGluAargPheCysGlnAargValSerPheLeuAsnCysSerLeu 142  
 QY 301 GACAAACGCGCTGCAGCATTTACTGCTAGAGAGTGCGCTGGCGGCGCTGTAGCTGT 360  
 DB 143 AspasnGlyGlyCysThrHisTyrcysLeuGluValGlyTrpAargCysSerCys 162  
 QY 361 GCGCTGGCTACAGCTGGGGGACGACCTCTGCTAGTGTACACCGCGAGTGAAGTTCCT 420  
 DB 163 AlaProGlyTrpLysLeuGlyAspLeuGluGlnCysHisProAlaValLysPhePro 182  
 QY 421 TGTGGAGCGCTGGAGCGGATGGAGAGAACGCGAGTCACCTGAAACGAGACAGAA 480  
 DB 183 CysGlyArgProTrpLysArgMetGluLysArgSerHisLeuLysArgAspThrGlu 202  
 QY 481 GACCAAGAACCAAGTAGATCCGCGCTCATTCATGGGAGATGACAGCGGGGAGAGAC 540

DB 203 AspGlnGluAspGlnValaspProArgLeuileaspGlyLysMetThrArgArgGlyasp 222  
 QY 541 AGCCCTGGCAGGTGGTCTCTGCTGACTCAAGAGAGAGCTGGCTGGGGGGAGTGCCTC 600  
 DB 223 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 242  
 QY 601 ATCCACCCCTCTCGGTGCTGACAGCGCCCTGATGATGATGATGATGATGATGATGATGAT 660  
 DB 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetaspGluSerLysLysLeuLeu 262  
 QY 661 GTCAGGCTGCGAGAGTATGACCTCGGGCGTGGAGAGTGGAGTGGAGTGGAGTGGAGTGC 720  
 DB 263 ValArgLeuGlyGluTyraaspLeuArgArgTrpLysTrpGluLeuAspLeuAspIle 282  
 QY 721 AAGAGGTCTCTGCTCCACCCCACTACAGCAAGAGACACCGACCAATGATGATGATGATGAT 780  
 DB 283 LysGluValPheValHisProAsnTyraaspSerThrThrAspAsnAspIleAlaLeu 302  
 QY 781 CTGACCTGCGCCAGCCCGCCCTCTCGCAGACCATAGTGCCTCTCTCTCTCTCTCTCTCT 840  
 DB 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrileValProIleCysLeuProasp 322  
 QY 841 AGCGGCTTTCAGAGCGGAGCTCAATCAGCGCGGCGGAGACCTCTGTCAGCGGCTGG 900  
 DB 323 SerGlyLeuAlaGluAargGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
 QY 901 GGCTACACACAGCAGCCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 DB 343 GlyThrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
 QY 961 ATCAAGATTCCTGCTGCTCGGCGACATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 DB 363 IleLysileProValProValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
 QY 1021 GAGAACATGCTGTGCTGGGCTATCTCGGGGACCGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
 DB 383 GluAsnMetLeuGlyAlaGlyileLeuGlyAspArgGlnAspAlaCysGluGlyaspSer 402  
 QY 1081 GGGGGGCGCATGCTGCTGCT 1140  
 DB 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
 QY 1141 GGTGAGGCTGTGGGCT 1200  
 DB 423 GlyGluGlyCysGlyLeuLeuHisAsnTyraaspGlyValTyraaspValSerArgTyraLeu 442  
 QY 1201 GACTGATCCATGGGCGACATCAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245  
 DB 443 AspTrpIleHisGlyHisileArgaspLysGluAlaProGlnLys 457

## RESULT 2

ID PRTC\_RABIT STANDARD; PRT; 458 AA.  
 AC Q28661;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Vitamin-K dependent protein C precursor (EC 3.4.21.69)  
 DE (Autoproteolysis IIA) (Anticoagulant protein C) (Blood coagulation  
 DE factor XIV) (Fragment).  
 OS PROC.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Shen L., He X., Dahlback B.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 CC !- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT  
 CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA

CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.  
CC -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va  
CC and VIIa.  
CC -!- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED  
CC INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE  
CC BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A  
CC TETRADECAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN: THIS  
CC REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS  
CC STRONGLY PROMOTED BY THROMBOMODULIN.  
CC -!- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.  
CC -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME  
CC GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.  
CC -!- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO  
CC ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING  
CC SITE IS NECESSARY FOR THE RECOGNITION OF THE  
CC THROMBIN-THROMBOMODULIN COMPLEX.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U49933; AA92956.1; -.  
CC HSSP: P04070; LPCU.  
CC DR MEROPS: S01.218; -.  
CC DR InterPro: IPR000152; Asx\_hydroxyl.  
CC DR InterPro: IPR000561; EGF-like.  
CC DR InterPro: IPR001881; EGF\_Ca.  
CC DR InterPro: IPR001254; Ser\_protease\_Try.  
CC DR InterPro: IPR000294; VitK\_dep\_GLA.  
CC DR Pfam: PF00008; EGF; 2.  
CC DR Pfam: PF00089; trypsin; 1.  
CC DR Pfam: PF00594; gla; 1.  
CC DR SMART: SM00181; EGF; 2.  
CC DR SMART: SM00069; GLA; 1.  
CC DR SMART: SM00020; TRYP-Spc; 1.  
CC DR PROSITE: PS00010; ASX-HYDROXYL; 1.  
CC DR PROSITE: PS00022; EGF\_1; 1.  
CC DR PROSITE: PS01186; EGF\_2; 2.  
CC DR PROSITE: PS01187; EGF\_CA; 1.  
CC DR PROSITE: PS00011; GLU CARBOXYLATION; 1.  
CC DR PROSITE: PS02040; TRYPSIN\_DOM; 1.  
CC DR PROSITE: PS00134; TRYPSIN\_SER; 1.  
CC DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
CC DR Blood coagulation; Glycoprotein; Serine protease;  
CC Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;  
CC EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.  
CC NON\_TER 1 1  
CC SIGNAL <1 27 BY SIMILARITY.  
CC PROPEP 28 36 BY SIMILARITY.  
CC CHAIN 37 458 VITAMIN K-DEPENDENT PROTEIN C.  
CC CHAIN 37 192 PROTEIN C LIGHT CHAIN (BY SIMILARITY).  
CC CHAIN 195 458 PROTEIN C HEAVY CHAIN (BY SIMILARITY).  
CC PEPTIDE 195 209 ACTIVATION PEPTIDE (BY SIMILARITY).  
CC SITE 209 210 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).  
CC DOMAIN 91 126 EGF-LIKE 1.  
CC DOMAIN 130 170 EGF-LIKE 2.  
CC DOMAIN 210 458 SERINE PROTEASE.  
CC MOD\_RES 42 42 GAMMA-CARBOXYGLUTAMIC ACID  
CC (BY SIMILARITY).  
CC MOD\_RES 43 43 GAMMA-CARBOXYGLUTAMIC ACID  
CC (BY SIMILARITY).  
CC MOD\_RES 50 50 GAMMA-CARBOXYGLUTAMIC ACID  
CC (BY SIMILARITY).  
CC MOD\_RES 52 52 GAMMA-CARBOXYGLUTAMIC ACID  
CC (BY SIMILARITY).  
CC MOD\_RES 55 55 GAMMA-CARBOXYGLUTAMIC ACID  
CC (BY SIMILARITY).

FT	MOD_RES	56	56	GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).
FT	MOD_RES	61	61	GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).
FT	MOD_RES	62	62	GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).
FT	MOD_RES	65	65	GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).
FT	MOD_RES	107	107	HYDROXYLATION (BY SIMILARITY).
FT	ACT_SITE	250	250	CHARGE RELAY SYSTEM.
FT	ACT_SITE	296	296	CHARGE RELAY SYSTEM.
FT	ACT_SITE	399	399	CHARGE RELAY SYSTEM.
FT	DISULFID	53	58	BY SIMILARITY.
FT	DISULFID	86	105	BY SIMILARITY.
FT	DISULFID	95	100	BY SIMILARITY.
FT	DISULFID	99	114	BY SIMILARITY.
FT	DISULFID	116	125	BY SIMILARITY.
FT	DISULFID	134	145	BY SIMILARITY.
FT	DISULFID	141	154	BY SIMILARITY.
FT	DISULFID	156	169	BY SIMILARITY.
FT	DISULFID	177	316	INTERCHAIN (BY SIMILARITY).
FT	DISULFID	235	251	BY SIMILARITY.
FT	DISULFID	370	384	BY SIMILARITY.
FT	DISULFID	395	423	BY SIMILARITY.
FT	CARBOHYD	133	133	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	287	287	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	352	352	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	458 AA;	51087 MW;	D75A5F990C8F29D7 CRC64;

Alignment Scores:

Pred. No.:	4.08e-92	Length:	458
Score:	1872.50	Matches:	336
Percent Similarity:	88.67%	Conservative:	32
Best Local Similarity:	80.96%	Mismatches:	44
Query Match:	80.12%	Indels:	3
DB:	1	Gaps:	2

US-09-763-153-1 (1-1245) x PRTC\_RABIT (1-458)

QY	1	GCCAACTCCTTCCTGGAGGAGTCCTCCGACACGACCGCTGGAGCGGAGTGCATAGAGAG	60
DB	37	AlaAsnSerPheLeuGluGluLeuArgProSerSerLeuGluArgGluCysValGluGlu	56
QY	61	ATCTGTGACTTCGAGGAGGCCAAGAAATTTTCCAAAATGTGGATGACACACTGCCTTC	120
DB	57	ValCysAspLeuGluGluAlaLysGluIlePheGlnSerValAspSerThrLeuAlaPhe	76
QY	121	TGGTCCAAAGCACGTCGACGGTGACACAGTCTTGGTCTTTCCTTGGACACCCGCTGC	180
DB	77	TrpTyrLysTyrValAspGlyAspGlnCysAlaAlaLeuProSerGluHisProCysSer	96
QY	181	AGCCTGTGCTCGGGCACGGCAGTCATCGACGGCATCGGCGCTTCAGCTTCAGCTGC	240
DB	97	SerGlnCysCysGlyHisGlyThrCysAlaAspSerIleGlyGlyPheSerCysGlnCys	116
QY	241	CCACGCGCTGGAGGCGCGCTTCCTCCGACGCGGAGTGCCTCCTCAATTGCTCGCTG	300
DB	117	HisGlyGlyTrpGluGlySerPheCysGlnTyrGluValArgPheSerAsnCysSerVal	136
QY	301	GACACGCGCTGCACGCATCTACTCTCCCTAGAGAGGTGGCGTGGCGGCTGTAGCTGT	360
DB	137	AspAsnGlyGlyCysAlaHisTyrCysLeuGluGluAlaGlyArgSerCysSerCys	156
QY	361	GGCCTGTGCTACAACTGGGGGACACCTCTCTGCTGCTACCGCGCTCAAGTTCCTC	420
DB	157	AlaProGlyTyrGluLeuAlaAspAspHisLeuGlnCysGluProAlaValArgPhePro	176
QY	421	TCTGGGAGG---CCCTGGAAGCGGATGGAGAAGACCGCAGTCACCTGAACAGCAGACA	477
DB	177	CysGlyArgLeuGlyTrpLysArgIleGluLysLysArgGlyAsnValLysArgAspLeu	196
QY	478	GAA-----GACCAAGAGACCAAGTAGATCCGGGCTCATTTGATGGGAAGATGACGAGG	531



FT	214	459	SERINE PROTEASE.
DOMAIN	214	459	GAMMA-CARBOXYGLUTAMIC ACID (BY
MOD_RES	47	47	SIMILARITY).
FT			GAMMA-CARBOXYGLUTAMIC ACID (BY
MOD_RES	48	48	SIMILARITY).
FT			GAMMA-CARBOXYGLUTAMIC ACID (BY
MOD_RES	55	55	SIMILARITY).
FT			GAMMA-CARBOXYGLUTAMIC ACID (BY
MOD_RES	57	57	SIMILARITY).
FT			GAMMA-CARBOXYGLUTAMIC ACID (BY
MOD_RES	60	60	SIMILARITY).
FT			GAMMA-CARBOXYGLUTAMIC ACID (BY
MOD_RES	61	61	SIMILARITY).
FT			GAMMA-CARBOXYGLUTAMIC ACID (BY
MOD_RES	66	66	SIMILARITY).
FT			GAMMA-CARBOXYGLUTAMIC ACID (BY
MOD_RES	67	67	SIMILARITY).
FT			GAMMA-CARBOXYGLUTAMIC ACID (BY
MOD_RES	70	70	SIMILARITY).
FT			HYDROXYLATION (BY SIMILARITY).
MOD_RES	112	112	CHARGE RELAY SYSTEM.
ACT_SITE	255	255	CHARGE RELAY SYSTEM.
ACT_SITE	301	301	CHARGE RELAY SYSTEM.
ACT_SITE	400	400	CHARGE RELAY SYSTEM.
DISULFID	58	63	BY SIMILARITY.
DISULFID	91	110	BY SIMILARITY.
DISULFID	100	105	BY SIMILARITY.
DISULFID	104	119	BY SIMILARITY.
DISULFID	121	130	BY SIMILARITY.
DISULFID	139	150	BY SIMILARITY.
DISULFID	146	159	BY SIMILARITY.
DISULFID	161	174	BY SIMILARITY.
DISULFID	182	321	INTERCHAIN (BY SIMILARITY).
DISULFID	240	256	BY SIMILARITY.
DISULFID	371	385	BY SIMILARITY.
DISULFID	396	424	BY SIMILARITY.
CARBOHYD	138	138	N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD	292	292	N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD	353	353	N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQUENCE	459 AA;	51866 MW;	8541AAC14CC16D09 CRC64;
Alignment Scores:			
Pred. No.:	1..12e-85	Length:	459
Score:	1750.50	Matches:	316
Percent Similarity:	84.21%	Conservative:	36
Best Local Similarity:	75.60%	Mismatches:	59
Query Match:	74.90%	Indels:	7
DB:	1	Gaps:	2
US-09-763-153-1 (1-1245) x PRTC_PIG (1-459)			
Qy	1	GCCAACTCTCTCTCGGAGGAGCTCCGTCACAGCAGCTGGAGCGGAGTGCATAGAGGAG	60
	42	AlaAsnSerPheLeuGluGluLeuArgProSerSerLeuGluArgGluCysLysGluGlu	61
	61	ATCTGTGACTTCGAGGAGGCCAAGGAAATTTCCAAATATGTGGATGCACACTGCCTTC	120
	62	ThrCysAspPheGluGluAlaArgGluIlePheGlnAsnThrGluAsnThrMetAlaPhe	81
Db			
	121	TGCTCCAAAGCACGTCGACGCTGACCACTGCTGGTCTCCGCTTGAGCACCCGCTGCCGC	180
Qy			
	82	TrpSerIlyThrHisaspGlyaspGlnCysAlaValSerProProGluHisLeucCysasp	101
Db			
	181	AGCCTGTGTCGGGGCAGCGACGTCAGTCGAGCGGATCGCGGACGCTTACGTGCGACATGC	240
Qy			
	102	SerProCysCysGlyargGlyThrCysIleAspGlyLeuGlyPheargCysaspCys	121
Db			
	241	CGCAGCGGCTGGGAGSGCCGCTCTCCGACGGCGAGGTGAGTCTCTCAATTGCTCGCTG	300
Qy			
	122	AlaGlnGlyTrpGluGlyArgPheCysLeuHisGluValArgPheSerAsnCysSerThr	141
Db			
	301	GACAACGGCGGCTGCACGCAATTACTGCCTAGAGGAGGTGGGTGGCGCGCTGTAGCTGT	360
Qy			
	142	GluAsnGlvGlvCysAlaHisTyrCysLeuGluGluGluGlvArgCysAlaCys	161
Db			

Alignment Scores:

Alignment Scores:	1.12e-85	Length:	459
Pred. No.:	1750.50	Matches:	316
Score:	84.21%	Conservative:	36
Percent Similarity:	75.60%	Mismatches:	59
Best Local Similarity:	74.90%	Indels:	7
Query Match:	1	Gaps:	2
DB:			

US-09-763-153-1 (1-1245) x PRTC\_PIG (1-459)

Qy	1	GCCAACTCCTTCCTGTGGAGGAGTCCGTACAGCAGCGCTGGAGCGGAGTGTCATAGAGGAG	60
	42	AlaAsnSerPheLeuGlulGluLeuArgProSerSerLeuGluArgGluCysLysGluGlu	61
	61	ATCTGTACTTCGAGGAGGCCAAGGAAATTTTCCAAAATGTGGATGCACACATGCGCCTTC	120
	62	ThrCysAspPheGluGluLaArgGluIlePheGlnAsnThrGluAsnThrMetAlaPhe	81
Db			
	121	TGTCCAAGCACATCGACGGTGACCAGTGCCTTGGCTCTGCCCTTGGAGCACCGCTGCGCCC	180
Qy			
	82	TrpSerLysTrHisAspGlyAspGlnCysAlaValSerProGluHisLeucCysAsp	101
Db			
	181	AGCTGTGTCGGGCACGCGCATGTCAGCGGATCGGCAGGTTCACGTGCGCACATGC	240
Qy			
	102	SerProcysCysGlyArgGlyThrCysIleAspGlyLeuGlyGlyPheArgCysaspCys	121
Db			
	241	CGCAGCGGCTGGGAGGCGCGCTTCTGCCACGCGAGGTGAGTTCCTCAATTGCTCGGTG	300
Qy			
	122	AlaGlnGlyTrpGluGlyArgPheCysLeuHisGluValArgPheSerAsnCysSerThr	141
Db			
	301	GACAACGGCGGCTGCACGCATTACTGCCTAGAGGAGGTGGGTGGCGGCGCTGTAGCTGT	360
Qy			
	142	GluAsnGlvGlcysAlaHisTrpCysLeuGluGluGluGlyGlyValArgCysAlaCys	161
Db			

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Bovinae; Bos.  
 CC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85014826; PubMed=6091100;  
 RA Long G.L., Balagaje R.M., McGillivray R.T.A.;  
 RL "Cloning and sequencing of liver cDNA coding for bovine protein C.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:5653-5656(1984).  
 RN [2]  
 RP SEQUENCE OF 40-194.  
 RX MEDLINE=83007325; PubMed=6896876;  
 RA Fernlund J., Stenflo J.;  
 RL "Amino acid sequence of the light chain of bovine protein C.";  
 RL J. Biol. Chem. 257:12170-12179(1982).  
 RN [3]  
 RP REVISION TO 110.  
 RX MEDLINE=83169769; PubMed=6572939;  
 RA Drakenberg T., Fernlund P., Roepstorff P., Stenflo J.;  
 RL "Beta-hydroxyaspartic acid in vitamin K-dependent protein C.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:1802-1806(1983).  
 RN [4]  
 RP SEQUENCE OF 197-456.  
 RX MEDLINE=83007326; PubMed=6896877;  
 RA Fernlund J., Fernlund P.;  
 RL "Amino acid sequence of the heavy chain of bovine protein C.";  
 RL J. Biol. Chem. 257:12180-12190(1982).  
 RN [5]  
 RP PROCESSING, AND CALCIUM-BINDING DATA.  
 RX MEDLINE=83213513; PubMed=6304092;  
 RA Esmon N.L., Debault L.E., Esmon C.T.;  
 RL "Proteolytic formation and properties of gamma-carboxyglutamic acid-  
 domainless protein C.";  
 RL J. Biol. Chem. 258:5548-5553(1983).  
 RN [6]  
 RP PROCESSING, AND CALCIUM-BINDING DATA.  
 RX MEDLINE=83213514; PubMed=6406503;  
 RA Johnson A.E., Esmon N.L., Laue T.M., Esmon C.T.;  
 RL "Structural changes required for activation of protein C are induced  
 by Ca<sup>2+</sup> binding to a high affinity site that does not contain gamma-  
 carboxyglutamic acid";  
 RL J. Biol. Chem. 258:5554-5560(1983).  
 CC -!- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT  
 REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA  
 IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.  
 CC -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va  
 and VIIIA.  
 CC -!- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED  
 INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE  
 BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A  
 TETRAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN; THIS  
 REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS  
 STRONGLY PROMOTED BY THROMBOMODULIN.  
 CC -!- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.  
 CC -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME  
 GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.  
 CC -!- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO  
 ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING  
 SITE IS NECESSARY FOR THE RECOGNITION OF THE  
 THROMBIN-THROMBOMODULIN COMPLEX.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; K02435; AAA30685.1; -.

DR PIR; A00928; KXBO.  
 DR HSP; P04070; LPCU.  
 DR MEROPS; S01.218; -.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR InterPro; IPR000294; Vitk\_dep\_GLA.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF00089; trypsin; 1.  
 DR Pfam; PF00594; gla; 1.  
 DR SMART; SM00181; EGF; 2.  
 DR SMART; SM00065; GLA; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 2.  
 DR PROSITE; PS01187; EGF\_Ca; 1.  
 DR PROSITE; PS00011; GLU CARBOXYLATION; 1.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; FALSE\_NEG.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Blood coagulation; Glycoprotein; Serine protease;  
 KW Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;  
 KW EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.  
 FT NON\_TER 1  
 FT SIGNAL <1 29  
 FT PROPEP 30 39  
 FT CHAIN 40 194  
 FT CHAIN 197 456  
 FT PEPTIDE 197 210  
 FT DOMAIN 94 129  
 FT DOMAIN 133 173  
 FT DOMAIN 211 456  
 FT MOD\_RES 45 45  
 FT MOD\_RES 46 46  
 FT MOD\_RES 53 53  
 FT MOD\_RES 55 55  
 FT MOD\_RES 58 58  
 FT MOD\_RES 59 59  
 FT MOD\_RES 62 62  
 FT MOD\_RES 64 64  
 FT MOD\_RES 65 65  
 FT MOD\_RES 68 68  
 FT MOD\_RES 74 74  
 FT MOD\_RES 110 110  
 FT ACT\_SITE 252 252  
 FT ACT\_SITE 298 298  
 FT ACT\_SITE 397 397  
 FT DISULFID 56 61  
 FT DISULFID 89 108  
 FT DISULFID 98 103  
 FT DISULFID 102 117  
 FT DISULFID 119 128  
 FT DISULFID 137 148  
 FT DISULFID 144 157  
 FT DISULFID 159 172  
 FT DISULFID 180 318  
 FT DISULFID 237 253  
 FT DISULFID 368 382  
 FT DISULFID 393 421  
 FT CARBOHYD 136 136  
 FT CARBOHYD 289 289  
 FT CARBOHYD 350 350  
 FT CARBOHYD 366 366  
 FT VARIANT 82 82  
 FT CONFLICT 455 456  
 SQ SEQUENCE 456 AA; 51407 MW; CAAP6833F894C209 CRC64;  
 VP -> PV (IN REF. 4).  
 N-LINKED (GLCNAC. . .)  
 N-LINKED (GLCNAC. . .)  
 N-LINKED (GLCNAC. . .)  
 F -> K.  
 CHARGE RELAY SYSTEM.  
 CHARGE RELAY SYSTEM.  
 CHARGE RELAY SYSTEM.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 INTERCHAIN.  
 Length: 456  
 Matches: 298  
 Conservative: 38  
 Alignment Scores:  
 Pred. No.: 6.74e-81  
 Score: 1660.00  
 Percent Similarity: 80.96%

Best Local Similarity: 71.81% Mismatches: 73  
 Query Match: 71.03% Indels: 6  
 DB: 1 Gaps: 2

US-09-763-153-1 (1-1245) x PRTC\_BOVIN (1-456)

QY 1 GCCAACTCTCTCTCGAGGAGTCCCTGCACAGCGCTCGAGCGGAGTGATAGAGGAG 60  
 Db 40 ALaAsnSerPheLeuGluGluLeuArgProGlyAsnValGluArgGluCysSerGluGlu 59  
 QY 61 ATCTGTCACTTCGAGGAGGCCAAGAAATTTCCAAATGTGGATGACACACTGGCTTC 120  
 Db 60 ValCysGluPheGluGluAlaArgGluLeuPheGlnAsnThrGluAspThrMetAlaPhe 79  
 QY 121 TGTCTCAGACGCTGACGGTGACAGTGTCTGCTGCTTCCCTTGGAGCACCGCTCGGCC 180  
 Db 80 TrpSerPheTyrSerAspGlyAspGlnCysGluAspArgProSerGlySerProCysAsp 99  
 QY 181 AGCTGTGCTGCGGCGACGCGCTGCATCGACGCGATCGGCGATTCAGCTGCGACTGC 240  
 Db 100 LeuProCysCysGlyArgGlyCysCysIleAspGlyLeuGlyGlyPheArgCysAspCys 119  
 QY 241 CGAGCGGCTGGAGGCGCGCTTTCGCCAGCGGAGGTGAGCTTCCTCAATTGCTGCTG 300  
 Db 120 AlaGluGlyTrpGluGlyArgPheCysLeuHisGluValArgPheSerAsnCysSerAla 139  
 QY 301 GACAAACGCGCTGACGCGCTTACTGCTAGAGAGGTGGCTGCGGCGCTGTAGCTGT 360  
 Db 140 GluAsnGlyGlyCysAlaHisTyrCysMetGluGluGlyArgHisCysSerCys 159  
 QY 361 GCGCTGCTGCTACAGCTGGGCGGAGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 Db 160 AlaProGlyTyrArgLeuGluAspAspHisGlnLeuCysValSerLysValThrPhePro 179  
 QY 421 TGTGGAGGCGCTGGAGCGGATGGAGAGAGCGGAGTCCAGTCAAGAGAGAGAGAGAG 480  
 Db 180 CysGlyArgLeuGlyLysArgMetGluLysLysLysLysLysLysLysLysLysLys 199  
 QY 481 -----GACCAAGAGACCAAGTAGATCCGCGCTCATGATGGGAGAGATGACAGCGG 534  
 Db 200 GlnValAspGlnLysAspGlnLeuAspProArgIleValAspGlyGlnGluAlaGlyTrp 219  
 QY 535 GGACAGCGCGCTGCGAGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594  
 Db 220 GlyGluSerProTrpGlnAlaValLeuLeuAspSerLysLysLysLysLysLysLys 239  
 QY 595 GTGCTCATCCACCGCTCTCGGTGCTGACAGCGCGCGCTGCTGCTGCTGCTGCTGCTG 654  
 Db 240 ValLeuLeuHisValSerTrpValLeuThrValAlaHisCysLysLysLysLysLys 259  
 QY 655 CTCCTGTGAGGCTGAGAGTATGAGCTGCGGCGCTGGGAGAGTGGGAGCTGGAGCTG 714  
 Db 260 LeuIleValArgLeuGlyGlyLysAspMetArgArgTrpGluSerTrpGluValAspLeu 279  
 QY 715 GACATCAAGAGGCTTCGCTCCACCCCACTACAGCAAGAGAGAGAGAGAGAGAGAGAG 774  
 Db 280 AspIleLysGluValIleHisProAsnTyrThrLysSerThrSerAspAsnAspIle 299  
 QY 775 GCATGTGTGACCG 834  
 Db 300 AlaLeuLeuArgLeuAlaLysProAlaThrLeuSerGlnThrIleValProIleCysLeu 319  
 QY 835 CCGAGACGCGCGCTGAGAGCGGAGCTCAATCAGCGCGCGCGCGCGCGCGCGCGCGCG 894  
 Db 320 ProAspSerGlyLeuSerGlyArgLysLeuThrGlnValGlnGluThrValValThr 339  
 QY 895 GGCTGGGCTACACAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 954  
 Db 340 GlyTrpGlyTyrArgAsp-----GluThrLysArgAsnArgThrPheValLeu 355  
 QY 955 AACTTCATCAAGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014  
 Db 356 SerPheIleLysValProValValProValProTyrAsnAlaCysValHisAlaMetGluAsnLys 375

QY 1015 GTGCTGAGAACATGCTGTGTGCGGCATCTCGGCGACCGGACGAGATGCTCGAGGCG 1074  
 Db 376 IleSerGluAsnMetLeuCysAlaGlyIleLeuGlyAspProArgAspAlaCysGluGly 395  
 QY 1075 GACAGTGGGGGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1134  
 Db 396 AspSerGlyGlyProMetValThrPhePheArgGlyThrTrpPheLeuValIleGlyLeuVal 415  
 QY 1135 AGCTGGGTGAGGCTGTGGGCTCTTACAACTACGCGCTTACACCAAGTCCAGCGCG 1194  
 Db 416 SerTrpGlyGluGlyCysGlyArgLeuTyrAsnTyrGlyValTyrThrLysValSerArg 435  
 QY 1195 TACTTCGACTGCATCCATGCGGCATCAGACAGCAAGGAGGCGCGC 1239  
 Db 436 TyrLeuAspTrpIleTyrGlyHisIleLysAlaGlnGluAlaPro 450

## RESULT 5

PRTC\_RAT  
 ID PRTC\_RAT STANDARD; PRT; 461 AA.

AC P31394;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Vitamin-K dependent protein C precursor (EC 3.4.21.69)  
 DE (Autoproteolytic protein C) (Anticoagulant protein C) (Blood coagulation factor XIV).  
 GN PROC.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]

## SEQUENCE FROM N.A.

RC STRAIN=Wistar; TISSUE=Liver;  
 RX MEDLINE=92329550; PubMed=1627650;  
 RA Okafuji T., Maekawa K., Nawa K., Marumoto Y.;  
 RT "The cDNA cloning and mRNA expression of rat protein C";  
 RL Biochim. Biophys. Acta 1131:329-332(1992).  
 CC -!- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT  
 CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA  
 CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.  
 CC -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va  
 CC and VIIIA.  
 CC -!- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED  
 CC INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE  
 CC BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A  
 CC TETRADECAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN; THIS  
 CC REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS  
 CC STRONGLY PROMOTED BY THROMBOMODULIN.  
 CC -!- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.  
 CC -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME  
 CC GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.  
 CC -!- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO  
 CC ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING  
 CC SITE IS NECESSARY FOR THE RECOGNITION OF THE  
 CC THROMBIN-THROMBOMODULIN COMPLEX.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X64336; CAA45617.1; -.  
 CC PIR; S18994; S18994.  
 CC PIR; S24312; S24312.  
 CC HSP; P04070; IPCU.  
 CC MEROPS; S01.218; -.

DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001881; EGF\_Ca.  
 DR InterPro: IPR002383; GLA\_blood.  
 DR InterPro: IPR001284; Ser\_protease\_Try.  
 DR InterPro: IPR000294; VitK\_dep\_GLA.  
 DR Pfam: PF00008; EGF; 2.  
 DR Pfam: PF00089; trypsin; 1.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PRINTS: PR00001; GLABLOOD.  
 DR SMART: SM00179; EGF\_CA; 1.  
 DR SMART: SM00001; EGF\_like; 1.  
 DR SMART: SM00069; GLA; 1.  
 DR SMART: SM00020; Tryp\_Spc; 1.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 2.  
 DR PROSITE: PS01187; EGF\_CA; 1.  
 DR PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE: PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 DR Blood coagulation; Glycoprotein; Serine protease;  
 KW Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;  
 KW EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.  
 FT SIGNAL 1 32  
 FT PROPEP 33 41  
 FT CHAIN 42 196  
 FT CHAIN 199 461  
 FT PEPTIDE 199 212  
 FT SITE 212 213  
 FT DOMAIN 96 131  
 FT DOMAIN 135 175  
 FT DOMAIN 213 461  
 FT MOD\_RES 47 47  
 FT MOD\_RES 48 48  
 FT MOD\_RES 55 55  
 FT MOD\_RES 57 57  
 FT MOD\_RES 60 60  
 FT MOD\_RES 61 61  
 FT MOD\_RES 66 66  
 FT MOD\_RES 67 67  
 FT MOD\_RES 70 70  
 FT MOD\_RES 112 112  
 FT ACT\_SITE 254 254  
 FT ACT\_SITE 300 300  
 FT ACT\_SITE 402 402  
 FT DISULFID 58 63  
 FT DISULFID 91 110  
 FT DISULFID 100 105  
 FT DISULFID 104 119  
 FT DISULFID 121 130  
 FT DISULFID 139 150  
 FT DISULFID 146 159  
 FT DISULFID 161 174  
 FT DISULFID 182 320  
 FT DISULFID 239 255  
 FT DISULFID 373 387  
 FT DISULFID 398 426  
 FT CARBOHYD 215 291  
 FT CARBOHYD 291 355  
 FT CARBOHYD 355 355

SQ SEQUENCE 461 AA; 51912 MW; 884CF93664EDACD5 CRC64;  
 Alignment Scores:  
 Pred. No.: 2.14e-80 Length: 461  
 Score: 1650.50 Matches: 289  
 Percent Similarity: 82.73% Conservative: 56  
 Best Local Similarity: 69.30% Mismatches: 69  
 Query Match: 70.62% Indels: 3  
 DB: 1 Gaps: 2  
 US-09-763-153-1 (1-1245) x PRTC\_RAT (1-461)  
 QY 1 GCCAACTCTCTCTGGAGGAGCTCGTCCACAGCAGCCCTGGAGCGAGTGCATAGAGCAG 60  
 DB 42 AlaasnSerPheLeuGluValArgAlaGlySerLeuGluArgGluCysMetGluGlu 61  
 QY 61 ATCTGTGACTTGGAGGAGGCCAAGAAATTTTCCAAAATGTGGATGACACACTGGCCCTC 120  
 DB 62 IleCysAspPheGluAlaGlnGluIlePheGlnAsnValGluAspThrLeuAlaPhe 81  
 QY 121 TGGTCCAAAGCAGTCGACGGTGACAGTGCTTGGTCTTGGCCCTGGAGACACCCGTCGCC 180  
 DB 82 TrpIleLysTyrPheAspGlyAspGlnCysSerThrProProLeuAspHisGlnCysAsp 101  
 QY 181 AGCCTGTGCTGGGGCAGCGGACGTCATCGACGCATCGGCAGCTTCAGCTGCAGCTGCC 240  
 DB 102 SerProCysGlyHisGlyThrCysIleAspGlyLeuGlyPheSerCysSerCys 121  
 QY 241 CGCAGCGCTGGAGGCGGCTTCTGCCAGCGCGAGGTGAGTCTCCTCAATGTCTCGCTG 300  
 DB 122 AsplysGlyTrpGluGlyArgPheCysGlnGlnGluMetGlyPheGlnAspCysArgVal 141  
 QY 301 GACAAAGCGGCTGACGCATCTACTGCTAGAGAGGTGGCTGGCGCGCTAGCTGT 360  
 DB 142 LysAsnGlyGlyCysTyrHisTyrCysLeuGluGluThrArgGlyArgCysArgCys 161  
 QY 361 CGCCTGGCTACAAGCTGGGGAGGACCTCTCGAGTGTACCCCGCAGTGAAGTTCCT 420  
 DB 162 AlaProGlyTyrGluLeuAlaAspHisMethCysArgProThrValAsnPhePro 181  
 QY 421 TGTGGAGGCGCTGGAGCGGATGGAGGAAGCGGAGTACCTGACCTGAAACGA- 474  
 DB 182 CysGlyLysLeuTrpLysArgThrAspLysLysArgLysAsnPheLysArgAspIleasp 201  
 QY 475 ACAGAAGACCAAGAACCAAGTAGTCCGCGCTCATGTATGATGGAGAGATGACAGCGG 534  
 DB 202 ProGluAspGluGluLeuGluGlyProArgIleValAsnGlyThrLeuThrLysGln 221  
 QY 535 GGAGACAGCCCTGGCAGTGTCTCTGGAGTCAAGAGAGAGAGAGTGGCTGGCGGGCA 594  
 DB 222 GlyAspSerProTrpGlnAlaIleLeuLeuAspSerLysLysLysLeuAlaCysGlyGly 241  
 QY 595 GTGCTCATCCACCCTCTCTGGGTGCTGACAGCGGCCCTGATGATGATGATGATGATGAT 654  
 DB 242 ValLeuIleHisThrSerTrpValLeuThrAlaAlaHisCysLeuGluSerSerLysLys 261  
 QY 655 CTCCTTGTGAGCTGTGAGAGTATGACCTGGCGGCTGGGAGAGAGTGGAGCTGGAGCTG 714  
 DB 262 LeuThrValArgLeuGlyGluTyrAspLeuArgArgAspProTrpGluLeuAspLeu 281  
 QY 715 GACATCAAGGAGGTCTTCTGCCACCCCACTACAGAGAGAGAGAGAGAGAGAGAGAGAG 774  
 DB 282 AspIleLysGluValLeuValHisProAsnTyrThrArgSerAsnSerAspAspIle 301  
 QY 775 GCACTGTGCACTGGGCCAGCCCGCCCTCTCGAGACCATAGTGGCCCATCTGCTCCT 834  
 DB 302 AlaLeuLeuArgLeuSerGlnProAlaThrLeuSerLysThrIleValProIleCysLeu 321  
 QY 835 CCGGACAGCGGCTTGGAGAGCGGAGCTCAATCAGCGCCCGGAGAGAGAGAGAGAGAGAG 894  
 DB 322 ProAsnSerGlyLeuAlaGln- --GluLeuSerGlnAlaGlyGlnGluThrValThr 340  
 QY 895 GGCTGGGCTTACCACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 954









REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA  
IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.  
-1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va  
and VIIIA.  
-1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.  
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; D43754; BAA07811.1; -;  
HSSP; P04070; IPCU.  
DR MEROPS; S01.218; -;  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00089; trypsin; 1.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; PARTIAL.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Blood coagulation; Glycoprotein; Serine protease; Hydrolase.  
FT NON\_TER 1 1  
FT ACT\_SITE 26 26 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 129 129 CHARGE RELAY SYSTEM.  
FT DISULFID 100 114 BY SIMILARITY.  
FT DISULFID 125 153 BY SIMILARITY.  
FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT NON\_TER 161 161  
SQ SEQUENCE 161 AA; 17770 MW; 27D78F185B2FCC69 CRC64;  
  
Alignment Scores:  
Pred. No.: 2.47e-37 Length: 161  
Score: 835.00 Matches: 154  
Percent Similarity: 98.14% Conservative: 4  
Best Local Similarity: 95.65% Mismatches: 3  
Query Match: 35.73% Indels: 0  
Gaps: 1  
DB:  
  
US-09-763-153-1 (1-1245) x PRTC\_MACMU (1-161)  
  
QY 694 GAGAGTGGAGCTGACATCGACATCAAGGAGGTCTTCGTCACCCCACTACAGCAAG 753  
Db 1 GluLysTrpGluLeuAspLeuAspIleGluGluValPheIleHisProAsnTrpThrLys 20  
  
QY 754 AGCACACCGACAATGACATCGACTGCTGACCTGGCCCGCCAGCCGACCTCTCGCAG 813  
Db 21 SerThrThrAspAsnAspIleAlaLeuLeuArgLeuAlaGlnProAlaThrLeuSerGln 40  
  
814 ACATAGTCCCATCTGCTCCCGACAGCGCCCTTGACAGCGCGGAGCTCAATCAGGCC 873  
41 ThrIleValProIleCysLeuProAspSerglyLeuAlaGluArgGluThrGlnAla 60  
  
QY 874 GCGCAGGAGACCTCGTGACGGGCTGGGCTACACAGCAGCCGAGAGAGGCCCAAG 933  
Db 61 GlyGlnGluThrLeuValThrGlyTrpGlyTyrHisSerSerglyGluLysGluAlaLys 80  
  
QY 934 AGAAACCGCACCTTCTGCTCAACTCATCATCAAGATTCGCGGTCCCGCACATGAGTGC 993  
Db 81 ArgAsnArgThrPheIleLeuAsnPheIleLysIleProValValProArgAsnGluLys 100  
  
QY 994 ACCGAGGTATGACGACATGTGTCTGAGACATGCTGTGCGGGCATCTCGGGGAC 1053  
Db 101 SerGluValMetSerAsnMetValSerGluAsnMetLeuCysAlaGlyIleLeuGlyAsp 120  
  
QY 1054 CGCAGGATCGCTGGGAGGCGACAGTGGGGGGCCCATGCTGCTCCCTTCACGCGCAC 1113  
Db 121 ArgGlnAspAlaCysGluGlyAspSerglyGlyProMetValAlaSerPheHisglyThr 140

QY 1114 TGGTTCTCGTGGCGCTGCTGAGCTGGGCTGAGGGCTGGGCTCTTCACAACTACGGC 1173  
Db 141 TrpPheLeuValGlyLeuValSerTrpGlyGlyCysGlyLeuLeuHisAsnTrpGly 160  
  
QY 1174 GTT 1176  
Db 161 Val 161  
  
RESULT 8  
FA10\_RABIT STANDARD; PRT; 490 AA.  
ID FA10\_RABIT  
AC O19045;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).  
GN F10.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97256311; PubMed=9101642;  
RA Pendurthi U.R., Anderson K.D., James H.L.;  
RT "Characterization of a full-length cDNA for rabbit factor X.";  
RL Thromb. Res. 85:503-514(1997).  
CC -1- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that  
converts prothrombin to thrombin in the presence of factor Va,  
calcium and phospholipid during blood clotting.  
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-I-Thr and then  
Arg-I-Ile bonds in prothrombin to form thrombin.  
CC -1- SUBUNIT: THE TWO CHAINS ARE FORMED FROM A SINGLE-CHAIN PRECURSOR  
BY THE EXCISION OF TWO ARG RESIDUES AND ARE HELD TOGETHER BY 1 OR  
MORE DISULFIDE BONDS.  
CC -1- PTM: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME  
GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND  
CALCIUM (BY SIMILARITY).  
CC -1- PTM: N- AND O-GLYCOSYLATED (BY SIMILARITY).  
CC -1- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE  
INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY)  
(BY SIMILARITY).  
CC -1- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO  
ANOTHER SITE, BEYOND THE GLA DOMAIN.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; AF003200; AAB62542.1; -;  
HSSP; P00742; 1HCG.  
DR MEROPS; S01.216; -;  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR002383; GLA\_blood.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR InterPro; IPR000294; VitK\_dep\_GLA.  
DR Pfam; PF00008; EGF; 2  
DR Pfam; PF00089; trypsin; 1.  
DR Pfam; PF00594; gla; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00001; GLABLOOD.  
DR SMART; SM00179; EGF\_CA; 1.  
DR SMART; SM00001; EGF\_like; 1.

DR SMART; SM00069; GLA; 1.  
 DR SMART; SM00020; TRYP\_SPC; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 2.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 DR PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE; PS02040; TRYP\_SIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYP\_SIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYP\_SIN\_SER; 1.  
 KW Glycoprotein; Hydroxylase; Serine protease; Plasma; Blood coagulation;  
 KW Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;  
 KW Signal; Zymogen; EGF-like domain; Repeat.  
 SIGNAL 1 20 POTENTIAL.  
 PROPEP 21 40 BY SIMILARITY.  
 CHAIN 41 180 FACTOR X LIGHT CHAIN.  
 CHAIN 184 490 FACTOR X HEAVY CHAIN.  
 FT PROPEP 184 232 ACTIVATION PEPTIDE.  
 FT CHAIN 233 490 ACTIVATED FACTOR XA, HEAVY CHAIN.  
 FT DOMAIN 86 122 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 125 165 EGF-LIKE 2.  
 FT DOMAIN 233 490 SERINE PROTEASE.  
 FT MOD\_RES 46 46 GAMMA-CARBOXYGLUTAMIC ACID (BY  
 SIMILARITY).  
 FT MOD\_RES 47 47 GAMMA-CARBOXYGLUTAMIC ACID (BY  
 SIMILARITY).  
 FT MOD\_RES 54 54 GAMMA-CARBOXYGLUTAMIC ACID (BY  
 SIMILARITY).  
 FT MOD\_RES 56 56 GAMMA-CARBOXYGLUTAMIC ACID (BY  
 SIMILARITY).  
 FT MOD\_RES 59 59 GAMMA-CARBOXYGLUTAMIC ACID (BY  
 SIMILARITY).  
 FT MOD\_RES 60 60 GAMMA-CARBOXYGLUTAMIC ACID (BY  
 SIMILARITY).  
 FT MOD\_RES 65 65 GAMMA-CARBOXYGLUTAMIC ACID (BY  
 SIMILARITY).  
 FT MOD\_RES 66 66 GAMMA-CARBOXYGLUTAMIC ACID (BY  
 SIMILARITY).  
 FT MOD\_RES 69 69 GAMMA-CARBOXYGLUTAMIC ACID (BY  
 SIMILARITY).  
 FT MOD\_RES 72 72 GAMMA-CARBOXYGLUTAMIC ACID (BY  
 SIMILARITY).  
 FT MOD\_RES 75 75 GAMMA-CARBOXYGLUTAMIC ACID (BY  
 SIMILARITY).  
 FT MOD\_RES 79 79 GAMMA-CARBOXYGLUTAMIC ACID (BY  
 SIMILARITY).  
 MOD\_RES 103 103 HYDROXYLATION (BY SIMILARITY).  
 ACT\_SITE 274 274 CHARGE RELAY SYSTEM.  
 ACT\_SITE 320 320 CHARGE RELAY SYSTEM.  
 ACT\_SITE 417 417 CHARGE RELAY SYSTEM.  
 FT DISULFID 90 101 BY SIMILARITY.  
 FT DISULFID 95 110 BY SIMILARITY.  
 FT DISULFID 112 121 BY SIMILARITY.  
 FT DISULFID 129 140 BY SIMILARITY.  
 FT DISULFID 136 149 BY SIMILARITY.  
 FT DISULFID 151 164 BY SIMILARITY.  
 FT DISULFID 172 340 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 239 244 BY SIMILARITY.  
 FT DISULFID 259 275 BY SIMILARITY.  
 FT DISULFID 388 402 BY SIMILARITY.  
 FT DISULFID 413 441 BY SIMILARITY.  
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 490 AA; 53965 MW; 3A39FA85AF2A6D11 CRC64;

## Alignment Scores:

Pred. No.: 2.88e-36 Length: 490  
 Score: 814.50 Matches: 173  
 Percent Similarity: 53.71% Conservative: 66  
 Best Local Similarity: 38.88% Mismatches: 159  
 Query Match: 34.85% Indels: 47  
 DB: 1 Gaps: 9

US-09-763-153-1 (1-1245) x FA10\_RABIT (1-490)  
 QY 1 GCCAACTCCCTTCCTGGAGGAGCTCGTCACAGACCTCGAGCGGAGTGCATAGAGAG 60  
 DB 41 AlaAsnSerPheLeuGluGluLeuLysGlyAsnLeuGluArgGluCysMetGluGlu 60  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGGATGACACACTGGCCCTC 120  
 DB 61 AsnCysSerTyrGluGluAlaLeuGluValPheGluAspArgGluLysThrAsnGluPhe 80  
 QY 121 TGGTCCAAAGCAGTCGAGGTCACCACTGCTTGTGCTTGGCCCTTGGAGCACCCTGGCC 180  
 DB 81 TrpAsnLysTyrValAspGlyAspGlnCysGluSerAsnPro 94  
 QY 181 AGCTGTGCTCGGGGACGCGCATCGACGGCATCGGCGATCGGACCTTCAGCTCGGACTGC 240  
 DB 95 -----CysGlnAsnGlnGlyThrCysLysAspGlyLeuGlyMetTyrThrCysSerCys 112  
 QY 241 CGCAGCGGCTGGAGGCGCGCTTCTGCCAGCGGAGGTGAGCTTCTCAATTCGCTGCTG 300  
 DB 113 ValGluGlyTyrGluGlyGlnAspCysGlu---ProValThrArgLysLeuCysSerLeu 131  
 QY 301 GACACCGCGGCTCCAGCATTACTGCTAGAGAGGTGGCGGCGGCTGTAGCTGT 360  
 DB 132 AspAsnGlyGlyCysAspGlnPheCysLysGluGluAsnSerValLeuCysSerCys 151  
 QY 361 GCGCTGGCTACAAAGCTGGGGGACGACCTCTGCTAGTGTCCACCCGAGTGAAGTCCCT 420  
 DB 152 AlaSerGlyTyrThrLeuGlyAspAsnGlyLysSerCysIleSerThrGluLeuPhePro 171  
 QY 421 TGTGGGAGG-----CCCTGGAAGCGGATGGAGAAAGCGCATCAC----- 462  
 DB 172 CysGlyLysValThrLeuGlyArgTyrArgSerProAlaThrAsnSerSerGluGly 191  
 QY 463 -----CTGAACGAGACACAGAACACACAGAA----- 489  
 DB 192 ProProGluAlaProGlyProGluGlnGlnAspGlyAsnLeuThrAlaThrGluAsn 211  
 QY 490 -----GACCAAGTAGATCCG----- 504  
 DB 212 ProPheAsnLeuLeuAspSerProGluProProGluAspSerSerSerLeuVal 231  
 QY 505 CGCTCATTCATGGGAAGATGACCCAGCGGGGAGACAGCCCTGGCAGGTGTCCTGCTG 564  
 DB 232 ArgIleValGlyGlyGlnAspCysArgAspGlyGluCysProTrpGlnAlaLeuVal 251  
 QY 565 GACTCAAAGAAGAGCTGGCGCTCGGGGAGTGTCTCATCCACCCCTCTGGTGTGACA 624  
 DB 252 AsnGluGluAsnGluGlyPheCysGlyGlyThrIleLeuSerGluTyrHisValLeuThr 271  
 QY 625 GCGGCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 684  
 DB 272 AlaAlaHisCysLeuHisGlnAlaLysArgPheLysValArgValGlyAspArgAspThr 291  
 QY 685 CGCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 744  
 DB 292 GluHisGluGluGlyAsnGluGluThrHisGluValGluValValValLysHisAsnArg 311  
 QY 745 TACAGCAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 804  
 DB 312 PheValLysGluThrTyrAspPheAspIleAlaValLeuArgLeuLysThrProIleThr 331  
 QY 805 CTCTCGCAGACCATAGTGCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 864  
 DB 332 PheArgArgAsnValAlaProAlaCysLeuProGlnLysAspTrpAlaGluSerThrLeu 351  
 QY 865 AATCAGCGCGCGGACGAGACCTCTGAGCGGCTGGGCTGAGCAGGACGAGGAGGAGGAG 924  
 DB 352 ----MetAlaGlnLysThrGlyIleValSerGlyPheGly-----ArgThrHis 366  
 QY 925 GAGGCGCAAGAGAAACCGGACCTTCTGCTCAACTTCATCAAGATTCCTGCTCCGCGAC 984



CC MORE DISULFIDE BONDS.  
 CC -!- PTM: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME  
 CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND  
 CC CALCIUM.  
 CC -!- PTM: N- AND O-GLYCOSYLATION.  
 CC -!- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE  
 CC INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).  
 CC -!- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO  
 CC ANOTHER SITE, BEYOND THE GLA DOMAIN.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X00673; CAA25286.1; .  
 CC PIR; A00925; EXBO.  
 CC PDB; IAP0; 31-JAN-94.  
 CC PDB; ICCF; 31-MAY-94.  
 CC PDB; IWHF; 15-MAY-97.  
 CC PDB; IWHF; 15-MAY-97.  
 CC MEROPS; S01.216; .  
 CC GlycoSuiteDB; P00743; .  
 CC InterPro; IPR000152; Asx\_hydroxyl.  
 CC InterPro; IPR000134; Chymotrypsin.  
 CC InterPro; IPR000561; EGF-like.  
 CC InterPro; IPR000742; EGF-2.  
 CC InterPro; IPR001881; EGF\_Ca.  
 CC InterPro; IPR002383; GLA\_blood.  
 CC InterPro; IPR001254; Ser\_protease\_Try.  
 CC InterPro; IPR000294; VitK\_dep\_GLA.  
 CC Pfam; PF00089; trypsin; 1.  
 CC Pfam; PF00594; gla; 1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC PRINTS; PR00001; GLABLOOD.  
 CC SMART; SM00179; EGF\_CA; 1.  
 CC SMART; SM00001; EGF\_like; 1.  
 CC SMART; SM00069; GLA; 1.  
 CC SMART; SM00020; TRYD\_SPC; 1.  
 CC PROSITE; PS00010; ASX\_HYDROXYL; 1.  
 CC PROSITE; PS00022; EGF\_1; 1.  
 CC PROSITE; PS01186; EGF\_2; 2.  
 CC PROSITE; PS01187; EGF\_CA; 1.  
 CC PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
 CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
 CC Glycoprotein; Hydrolase; Serine protease; Blood coagulation;  
 CC Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;  
 CC Signal; Zymogen; EGF-like domain; Repeat; Sulfation; 3D-structure.  
 CC SIGNAL 1 23 POTENTIAL.  
 CC PROPEP 24 40  
 CC CHAIN 41 180 FACTOR X LIGHT CHAIN.  
 CC CHAIN 183 492 FACTOR X HEAVY CHAIN.  
 CC PROPEP 183 233 ACTIVATION PEPTIDE.  
 CC CHAIN 234 492 ACTIVATED FACTOR XA, HEAVY CHAIN.  
 CC PROPEP 476 492 MAY BE REMOVED BUT IS NOT NECESSARY FOR  
 CC ACTIVATION.  
 CC DOMAIN 86 122 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).  
 CC DOMAIN 125 165 EGF-LIKE 2.  
 CC DOMAIN 234 492 SERINE PROTEASE.  
 CC ACT\_SITE 275 275 CHARGE RELAY SYSTEM.  
 CC ACT\_SITE 321 321 CHARGE RELAY SYSTEM.  
 CC ACT\_SITE 418 418 CHARGE RELAY SYSTEM.  
 CC MOD\_RES 46 46 GAMMA-CARBOXYGLUTAMIC ACID.  
 CC MOD\_RES 47 47 GAMMA-CARBOXYGLUTAMIC ACID.  
 CC MOD\_RES 54 54 GAMMA-CARBOXYGLUTAMIC ACID.

FT	MOD_RES	56	56	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	59	59	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	60	60	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	65	65	GAMMA-CARBOXYGLUTAMIC ACID.
FT	MOD_RES	66	66	GAMMA-CARBOXYGLUTAMIC ACID.

Alignment Scores: 5.62e-36 Length: 492  
 Pred. No.: 809.00 Matches: 171  
 Score: 52.27% Conservative: 71  
 Percent Similarity: 36.93% Mismatches: 149  
 Best Local Similarity: 34.62% Indels: 72  
 Query Match: 1 Gaps: 12  
 DB: 1

US-09-763-153-1 (1-1245) x FA10\_BOVIN (1-492)

QY	1	GCCAACTCTCTCTGGAGGAGCTCCGTCACAGCAGCGCTGGAGCGGAGTGTCATAGAGGAG	60
DB	41	AlaAsnSerPheLeuGluGluValLysGlnGlyAsnLeuGluArgGluCysLeuGluGlu	60
QY	61	ATCTGTGACTTCGAGGAGGCAAGAAATTTCCAAATGTGGATGACACACTGGCCTTC	120
DB	61	AlaCysSerLeuGluGluAlaArgGluValPheGluAspAlaGluGlnThrAspGluPhe	80
QY	121	TGCTCCAAAGCAGCTCGACGCTGACAGTGTCTTGGTCTTGGCTTGGACACCGCTGGCC	180
DB	81	TrpSerLysTyrLysAspGlyAspGlnCys-----GluGlyHisProCysLeu	96
QY	181	AGCCTGTGCTGGCGGACGCGACGTCATCGACGCGATCGGACGTCAGCTTCAGCTCGACTGC	240
DB	97	Asn-----GlnGlyHisCysLysAspGlyGlyAspThrCysThrCys	112
QY	241	CGCAGCGCTGGAGGCGCTCTTCCAG-----CGCAGGTGAGCTTCTCAAT	291
DB	113	AlaGluGlyPheGluGlyLysAsnCysGluPheSerThrArgGluLeu	128
QY	292	TGCTCGCTGGACACGCGGCTGCACGCTTACTGCTAGAGGAGTGGCTGGCGGCGC	351
DB	129	CysSerLeuAspAsnGlyGlyCysAspGlnPheCysArgGluGluArgSerGluValArg	148
QY	352	TGTAGCTGTGCGCTGGCTACAAAGCTGGGGGACGACTCTCGAGTGTCCACCCGCGAGTG	411
DB	149	CysSerCysAlaHisGlyTyrValLeuGlyAspSerLysSerCysValSerThrGlu	168
QY	412	AAGTTCCTGTGGAGG-----CCCTGGAAGCGGATGGAAGAAGCGCAGCTACCTGAA	468
DB	169	ArgPheProCysGlyLysPheThrGlnGlyArgSerArgArgTrpAlaIleHisThrSer	188
QY	469	CGAGACACAGAACAC-----CAAGAAGACCAAGTAGATCCG-----	504
DB	189	GluAspAlaLeuAspAlaSerGluLeuGluHisTyrAspProAlaAspLeuSerProThr	208
QY	504	-----	504
DB	209	GluSerSerLeuAspLeuLeuGlyLeuAsnArgThrGluProSerAlaGlyGluAspGly	228
QY	505	-----CGGCTCATTTGATGGAGAGATGACCGGGGGGAGACAGCCCTGGGAG	552
DB	229	SerGlnValValArgIleValGlyArgAspCysAlaGluGlyGluCysProThrGln	248
QY	553	GTGGTCTCTGTGGACTCAAGAGAGAGCTGGCTGGGGGCGAGTGTCTATCCACCCCTCC	612
DB	249	AlaLeuLeuValAsnGluGluAsnGluGlyPheCysGlyGlyThrIleLeuAsnGluPhe	268
QY	613	TGGGTCTGACAGCGGCCACTGATGATGATGATGATGATGATGATGATGATGATGATG	672
DB	269	TyrValLeuThrAlaAlaHisCysLeuHisGlnAlaLysArgPheThrValArgValGly	288
QY	673	GAGTATGACCTGGCGGCTGGGAGAGAGTGGGAGCTGGACCTGACATCAAGAGAGTCTTC	732
DB	289	AspArgAsnThrGluGlnGluGluGlyAsnGluMetAlaHisGluValGluMetThrVal	308
QY	733	GTCCACCCCACTACAGCAGAGACCCAGCAGATGACATGACATGCTGCTGCTGACCTGGCC	792



Tue Mar 18 16:19:33 2003

FT	DISULFID	217	233	BY SIMILARITY.	646	TCCAGGAAGCTCTGTCAGGCTTGAGAGTATGACCTGGCGCTGGGAGAGTGGGAG	705
FT	DISULFID	349	368	BY SIMILARITY.	240	LeuArgAsnLeuThrIleValLeuGluHisAspLeuSerGluHisGluGlyAspGlu	259
FT	DISULFID	379	407	BY SIMILARITY.	706	CTGGACCTGGACATCAGAGGAGTCTTCGTCACCCCACTACAGCAAGAGCACCACCGAC	765
FT	MOD_RES	45	45	GAMMA-CARBOXYGLUTAMIC ACID.	260	GlnValArgHisValAlaGlnLeuIleMetProAspLysTyrValProGlyLysThrAsp	279
FT	MOD_RES	46	46	GAMMA-CARBOXYGLUTAMIC ACID.	766	AATGACATCGACATCGCTGACCTGGCCCGCCACCTCTCGCAGACCATAGTGGCC	825
FT	MOD_RES	53	53	GAMMA-CARBOXYGLUTAMIC ACID.	280	HisAspIleAlaLeuLeuArgLeuGlnProAlaLeuThrAsnAsnValValPro	299
FT	MOD_RES	55	55	GAMMA-CARBOXYGLUTAMIC ACID.	826	ATCTGCTCTCCCGGACGCGCTTCGACAGCGCGAGCTCAATCAGCCCGCCGAGGAC	885
FT	MOD_RES	58	58	GAMMA-CARBOXYGLUTAMIC ACID.	300	LeuCysLeuProGluArgAsnPheSerGluSerThrLeuAlaThrIle	318
FT	MOD_RES	64	64	GAMMA-CARBOXYGLUTAMIC ACID.	886	CTGCTGACCGGGTGGGGC-----TACACAGCAGCGGAGAGGAGGAGGAGAGAG	936
FT	MOD_RES	65	65	GAMMA-CARBOXYGLUTAMIC ACID.	319	ArgValSerGlyTrpGlyGlnLeuLeuTyrArgGlyAlaLeuAlaArgGlu	335
FT	MOD_RES	68	68	GAMMA-CARBOXYGLUTAMIC ACID.	937	AACGCGACCTTCCTCCTCACTTCAATTCAGATTCCTCGTGGTCCGCGACATGAGTGCAGC	996
FT	MOD_RES	74	74	GAMMA-CARBOXYGLUTAMIC ACID.	336	-----LeuMetAlaIleAspValProArgLeuMetThrGlnAspCysVal	350
FT	MOD_RES	102	102	HYDROXYLATION (BY SIMILARITY).	997	GAGTGCATG-----ACCAATGCTGTCTGAGACATGCTGTGTGGCGG	1041
FT	CARBOHYD	211	211	N-LINKED (GLCNAC. . .) (POTENTIAL).	351	GluGlnSerGluHisLysProGlySerProGluValThrGlyAsnMetPheCysAlaGly	370
FT	CARBOHYD	242	242	N-LINKED (GLCNAC. . .) (POTENTIAL).	1042	ATCTCTGGGGACCGGACGAGTGGTGGGCGGACAGTGGGGGGCCCATGTCGCTCC	1101
FT	CARBOHYD	306	306	N-LINKED (GLCNAC. . .) (POTENTIAL).	371	TyrLeuAspGlySerLysAspAlaCysLysGlyAspSerGlyGlyProHisAlaThrSer	390
FT	SEQUENCE	444	444	SEQUENCE	1102	TTCCACGCGCACCTGCTGCTGCTGGGCTGGTGGGCTGGGCTGGGCTGGGCTCCTT	1161
FT	SEQUENCE	444	444	SEQUENCE	391	TyrHisGlyThrTyrTyrLeuThrGlyValValSerTrpGlyGlyCysAlaVal	410
FT	SEQUENCE	444	444	SEQUENCE	1162	CACAACTACGCGGTTTACACAAAGTACGCGCTACCTCGATCGATCGATCGATCGATC	1221
FT	SEQUENCE	444	444	SEQUENCE	411	GlyHisValGlyValTyrThrArgValSerArgTyrThrGluTrpLeuSerArgLeuMet	430
FT	SEQUENCE	444	444	SEQUENCE	1222	AGAGCAAG 1230	
FT	SEQUENCE	444	444	SEQUENCE	431	ArgSerLys 433	
FT	SEQUENCE	444	444	SEQUENCE	FA10_CHICK	STANDARD;	PRT; 475 AA.
FT	SEQUENCE	444	444	SEQUENCE	ID	FA10_CHICK	
FT	SEQUENCE	444	444	SEQUENCE	AC	P25155;	
FT	SEQUENCE	444	444	SEQUENCE	DT	01-MAY-1992 (Rel. 22, Created)	
FT	SEQUENCE	444	444	SEQUENCE	DT	01-MAY-1992 (Rel. 22, Last sequence update)	
FT	SEQUENCE	444	444	SEQUENCE	DT	15-JUN-2002 (Rel. 41, Last annotation update)	
FT	SEQUENCE	444	444	SEQUENCE	DE	Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor)	
FT	SEQUENCE	444	444	SEQUENCE	DE	(Virus activating protease) (VAP).	
FT	SEQUENCE	444	444	SEQUENCE	GN	FX.	
FT	SEQUENCE	444	444	SEQUENCE	OS	Gallus gallus (Chicken).	
FT	SEQUENCE	444	444	SEQUENCE	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
FT	SEQUENCE	444	444	SEQUENCE	OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;	
FT	SEQUENCE	444	444	SEQUENCE	OC	Gallus.	
FT	SEQUENCE	444	444	SEQUENCE	OX	NCBI_TaxID=9031;	
FT	SEQUENCE	444	444	SEQUENCE	RN	[1]	
FT	SEQUENCE	444	444	SEQUENCE	RP	SEQUENCE FROM N.A.	
FT	SEQUENCE	444	444	SEQUENCE	RC	TISSUE=Chorioallantoic membrane;	
FT	SEQUENCE	444	444	SEQUENCE	RC	MEDLINE=91257322; PubMed=2044767;	
FT	SEQUENCE	444	444	SEQUENCE	RA	Suzuki H., Harada A., Hayashi Y., Wada K., Asaka J.-I., Gotoh B.,	
FT	SEQUENCE	444	444	SEQUENCE	RA	Ogasawara T., Nagai Y.	
FT	SEQUENCE	444	444	SEQUENCE	RT	"Primary structure of the virus activating protease from chick	
FT	SEQUENCE	444	444	SEQUENCE	RT	embryo. Its identity with the blood clotting factor Xa."	
FT	SEQUENCE	444	444	SEQUENCE	RL	FEBS Lett. 283:281-285(1991).	
FT	SEQUENCE	444	444	SEQUENCE	RN	[2]	
FT	SEQUENCE	444	444	SEQUENCE	RP	SEQUENCE OF 41-55 AND 241-261.	
FT	SEQUENCE	444	444	SEQUENCE	RC	TISSUE=Allantoic fluid;	
FT	SEQUENCE	444	444	SEQUENCE	RX	MEDLINE=91065352; PubMed=2174359;	
FT	SEQUENCE	444	444	SEQUENCE	RX	Gotoh B., Ogasawara T., Toyoda T., Inocencio N.M., Hamaguchi M.,	

FT	DISULFID	217	233	BY SIMILARITY.	646	TCCAGGAAGCTCTGTCAGGCTTGAGAGTATGACCTGGCGCTGGGAGAGTGGGAG	705
FT	DISULFID	349	368	BY SIMILARITY.	240	LeuArgAsnLeuThrIleValLeuGluHisAspLeuSerGluHisGluGlyAspGlu	259
FT	DISULFID	379	407	BY SIMILARITY.	706	CTGGACCTGGACATCAGAGGAGTCTTCGTCACCCCACTACAGCAAGAGCACCACCGAC	765
FT	MOD_RES	45	45	GAMMA-CARBOXYGLUTAMIC ACID.	260	GlnValArgHisValAlaGlnLeuIleMetProAspLysTyrValProGlyLysThrAsp	279
FT	MOD_RES	46	46	GAMMA-CARBOXYGLUTAMIC ACID.	766	AATGACATCGACATCGCTGACCTGGCCCGCCACCTCTCGCAGACCATAGTGGCC	825
FT	MOD_RES	53	53	GAMMA-CARBOXYGLUTAMIC ACID.	280	HisAspIleAlaLeuLeuArgLeuGlnProAlaLeuThrAsnAsnValValPro	299
FT	MOD_RES	55	55	GAMMA-CARBOXYGLUTAMIC ACID.	826	ATCTGCTCTCCCGGACGCGCTTCGACAGCGCGAGCTCAATCAGCCCGCCGAGGAC	885
FT	MOD_RES	58	58	GAMMA-CARBOXYGLUTAMIC ACID.	300	LeuCysLeuProGluArgAsnPheSerGluSerThrLeuAlaThrIle	318
FT	MOD_RES	64	64	GAMMA-CARBOXYGLUTAMIC ACID.	886	CTGCTGACCGGGTGGGGC-----TACACAGCAGCGGAGAGGAGGAGGAGAGAG	936
FT	MOD_RES	65	65	GAMMA-CARBOXYGLUTAMIC ACID.	319	ArgValSerGlyTrpGlyGlnLeuLeuTyrArgGlyAlaLeuAlaArgGlu	335
FT	MOD_RES	68	68	GAMMA-CARBOXYGLUTAMIC ACID.	937	AACGCGACCTTCCTCCTCACTTCAATTCAGATTCCTCGTGGTCCGCGACATGAGTGCAGC	996
FT	MOD_RES	74	74	GAMMA-CARBOXYGLUTAMIC ACID.	336	-----LeuMetAlaIleAspValProArgLeuMetThrGlnAspCysVal	350
FT	MOD_RES	102	102	HYDROXYLATION (BY SIMILARITY).	997	GAGTGCATG-----ACCAATGCTGTCTGAGACATGCTGTGTGGCGG	1041
FT	CARBOHYD	211	211	N-LINKED (GLCNAC. . .) (POTENTIAL).	351	GluGlnSerGluHisLysProGlySerProGluValThrGlyAsnMetPheCysAlaGly	370
FT	CARBOHYD	242	242	N-LINKED (GLCNAC. . .) (POTENTIAL).	1042	ATCTCTGGGGACCGGACGAGTGGTGGGCGGACAGTGGGGGGCCCATGTCGCTCC	1101
FT	CARBOHYD	306	306	N-LINKED (GLCNAC. . .) (POTENTIAL).	371	TyrLeuAspGlySerLysAspAlaCysLysGlyAspSerGlyGlyProHisAlaThrSer	390
FT	SEQUENCE	444	444	SEQUENCE	1102	TTCCACGCGCACCTGCTGCTGCTGGGCTGGTGGGCTGGGCTGGGCTGGGCTCCTT	1161
FT	SEQUENCE	444	444	SEQUENCE	391	TyrHisGlyThrTyrTyrLeuThrGlyValValSerTrpGlyGlyCysAlaVal	410
FT	SEQUENCE	444	444	SEQUENCE	1162	CACAACTACGCGGTTTACACAAAGTACGCGCTACCTCGATCGATCGATCGATC	1221
FT	SEQUENCE	444	444	SEQUENCE	411	GlyHisValGlyValTyrThrArgValSerArgTyrThrGluTrpLeuSerArgLeuMet	430
FT	SEQUENCE	444	444	SEQUENCE	1222	AGAGCAAG 1230	
FT	SEQUENCE	444	444	SEQUENCE	431	ArgSerLys 433	
FT	SEQUENCE	444	444	SEQUENCE	FA10_CHICK	STANDARD;	PRT; 475 AA.
FT	SEQUENCE	444	444	SEQUENCE	ID	FA10_CHICK	
FT	SEQUENCE	444	444	SEQUENCE	AC	P25155;	
FT	SEQUENCE	444	444	SEQUENCE	DT	01-MAY-1992 (Rel. 22, Created)	
FT	SEQUENCE	444	444	SEQUENCE	DT	01-MAY-1992 (Rel. 22, Last sequence update)	
FT	SEQUENCE	444	444	SEQUENCE	DT	15-JUN-2002 (Rel. 41, Last annotation update)	
FT	SEQUENCE	444	444	SEQUENCE	DE	Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor)	
FT	SEQUENCE	444	444	SEQUENCE	DE	(Virus activating protease) (VAP).	
FT	SEQUENCE	444	444	SEQUENCE	GN	FX.	
FT	SEQUENCE	444	444	SEQUENCE	OS	Gallus gallus (Chicken).	
FT	SEQUENCE	444	444	SEQUENCE	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
FT	SEQUENCE	444	444	SEQUENCE	OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;	
FT	SEQUENCE	444	444	SEQUENCE	OC	Gallus.	
FT	SEQUENCE	444	444	SEQUENCE	OX	NCBI_TaxID=9031;	
FT	SEQUENCE	444	444	SEQUENCE	RN	[1]	
FT	SEQUENCE	444	444	SEQUENCE	RP	SEQUENCE FROM N.A.	
FT	SEQUENCE	444	444	SEQUENCE	RC	TISSUE=Chorioallantoic membrane;	
FT	SEQUENCE	444	444	SEQUENCE	RC	MEDLINE=91257322; PubMed=2044767;	
FT	SEQUENCE	444	444	SEQUENCE	RA	Suzuki H., Harada A., Hayashi Y., Wada K., Asaka J.-I., Gotoh B.,	
FT	SEQUENCE	444	444	SEQUENCE	RA	Ogasawara T., Nagai Y.	
FT	SEQUENCE	444	444	SEQUENCE	RT	"Primary structure of the virus activating protease from chick	
FT	SEQUENCE	444	444	SEQUENCE	RT	embryo. Its identity with the blood clotting factor Xa."	
FT	SEQUENCE	444	444	SEQUENCE	RL	FEBS Lett. 283:281-285(1991).	
FT	SEQUENCE	444	444	SEQUENCE	RN	[2]	
FT	SEQUENCE	444	444	SEQUENCE	RP	SEQUENCE OF 41-55 AND 241-261.	
FT	SEQUENCE	444	444	SEQUENCE	RC	TISSUE=Allantoic fluid;	
FT	SEQUENCE	444	444	SEQUENCE	RX	MEDLINE=91065352; PubMed=2174359;	
FT	SEQUENCE	444	444	SEQUENCE	RX	Gotoh B., Ogasawara T., Toyoda T., Inocencio N.M., Hamaguchi M.,	

FT	DISULFID	217	233	BY SIMILARITY.	646	TCCAGGAAGCTCTGTCAGGCTTGAGAGTATGACCTGGCGCTGGGAGAGTGGGAG	705
FT	DISULFID	349	368	BY SIMILARITY.	240	LeuArgAsnLeuThrIleValLeuGluHisAspLeuSerGluHisGluGlyAspGlu	259
FT	DISULFID	379	407	BY SIMILARITY.	706	CTGGACCTGGACATCAGAGGAGTCTTCGTCACCCCACTACAGCAAGAGCACCACCGAC	765
FT	MOD_RES	45	45	GAMMA-CARBOXYGLUTAMIC ACID.	260	GlnValArgHisValAlaGlnLeuIleMetProAspLysTyrValProGlyLysThrAsp	279
FT	MOD_RES	46	46	GAMMA-CARBOXYGLUTAMIC ACID.	766	AATGACATCGACATCGCTGACCTGGCCCGCCACCTCTCGCAGACCATAGTGGCC	825
FT	MOD_RES	53	53	GAMMA-CARBOXYGLUTAMIC ACID.	280	HisAspIleAlaLeuLeuArgLeuGlnProAlaLeuThrAsnAsnValValPro	299
FT	MOD_RES	55	55	GAMMA-CARBOXYGLUTAMIC ACID.	826	ATCTGCTCTCCCGGACGCGCTTCGACAGCGCGAGCTCAATCAGCCCGCCGAGGAC	885
FT	MOD_RES	58	58	GAMMA-CARBOXYGLUTAMIC ACID.	300	LeuCysLeuProGluArgAsnPheSerGluSerThrLeuAlaThrIle	318
FT	MOD_RES	59	59	GAMMA-CARBOXYGLUTAMIC ACID.	886	CTGCTGACCGGGTGGGGC-----TACACAGCAGCGGAGAGGAGGAGGAGAGAG	936
FT	MOD_RES	64	64	GAMMA-CARBOXYGLUTAMIC ACID.	319	ArgValSerGlyTrpGlyGlnLeuLeuTyrArgGlyAlaLeuAlaArgGlu	335
FT	MOD_RES	65	65	GAMMA-CARBOXYGLUTAMIC ACID.	937	AACGCGACCTTCCTCCTCACTTCAATTCAGATTCCTCGTGGTCCGCGACATGAGTGCAGC	996
FT	MOD_RES	68	68	GAMMA-CARBOXYGLUTAMIC ACID.	336	-----LeuMetAlaIleAspValProArgLeuMetThrGlnAspCysVal	350
FT	MOD_RES	74	74	GAMMA-CARBOXYGLUTAMIC ACID.	997	GAGTGCATG-----ACCAATGCTGTCTGAGACATGCTGTGTGGCGG	1041
FT	MOD_RES	102	102	HYDROXYLATION (BY SIMILARITY).	351	GluGlnSerGluHisLysProGlySerProGluValThrGlyAsnMetPheCysAlaGly	370
FT	CARBOHYD	211	211	N-LINKED (GLCNAC. . .) (POTENTIAL).	371	ATCTCTGGGGACCGGACGAGTGGTGGGCGGACAGTGGGGGGCCCATGTCGCTCC	430
FT	CARBOHYD	242	242	N-LINKED (GLCNAC. . .) (POTENTIAL).	431	TyrLeuAspGlySerLysAspAlaCysLysGlyAspSerGlyGlyProHisAlaThrSer	490
FT	CARBOHYD	306	306	N-LINKED (GLCNAC. . .) (POTENTIAL).	491	TTCCACGCGCACCTGCTGCTGCTGGGCTGGTGGGCTGGGCTGGGCTCCTT	550
FT	SEQUENCE	444	444	SEQUENCE	551	TyrHisGlyThrTyrTyrLeuThrGlyValValSerTrpGlyGlyCysAlaVal	610
FT	SEQUENCE	444	444	SEQUENCE	611	CACAACTACGCGGTTTACACAAAGTACGCGCTACCTCGATCGATCGATC	670
FT	SEQUENCE	444	444	SEQUENCE	671	GlyHisValGlyValTyrThrArgValSerArgTyrThrGluTrpLeuSerArgLeuMet	730
FT	SEQUENCE	444	444	SEQUENCE	731	AGAGCAAG 740	
FT	SEQUENCE	444	444	SEQUENCE	741	ArgSerLys 743	
FT	SEQUENCE	444	444	SEQUENCE	FA10_CHICK	STANDARD;	PRT; 475 AA.
FT	SEQUENCE	444	444	SEQUENCE	ID	FA10_CHICK	
FT	SEQUENCE	444	444	SEQUENCE	AC	P25155;	
FT	SEQUENCE	444	444	SEQUENCE	DT	01-MAY-1992 (Rel. 22, Created)	
FT	SEQUENCE	444	444	SEQUENCE	DT	01-MAY-1992 (Rel. 22, Last sequence update)	
FT	SEQUENCE	444	444	SEQUENCE	DT	15-JUN-2002 (Rel. 41, Last annotation update)	
FT	SEQUENCE	444	444	SEQUENCE	DE	Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor)	
FT	SEQUENCE	444	444	SEQUENCE	DE	(Virus activating protease) (VAP).	
FT	SEQUENCE	444	444	SEQUENCE	GN	FX.	
FT	SEQUENCE	444	444	SEQUENCE	OS	Gallus gallus (Chicken).	
FT	SEQUENCE	444	444	SEQUENCE	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
FT	SEQUENCE	444	444	SEQUENCE	OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;	
FT	SEQUENCE	444	444	SEQUENCE	OC	Gallus.	
FT	SEQUENCE	444	444	SEQUENCE	OX	NCBI_TaxID=9031;	
FT	SEQUENCE	444	444	SEQUENCE	RN	[1]	
FT	SEQUENCE	444	444	SEQUENCE	RP	SEQUENCE FROM N.A.	
FT	SEQUENCE	444	444	SEQUENCE	RC	TISSUE=Chorioallantoic membrane;	
FT	SEQUENCE	444	444	SEQUENCE	RC	MEDLINE=91257322; PubMed=2044767;	
FT	SEQUENCE	444	444	SEQUENCE	RA	Suzuki H., Harada A., Hayashi Y., Wada K., Asaka J.-I., Gotoh B.,	
FT	SEQUENCE	444	444	SEQUENCE	RA	Ogasawara T., Nagai Y.	
FT	SEQUENCE	444	444	SEQUENCE	RT	"Primary structure of the virus activating protease from chick	
FT	SEQUENCE	444	444	SEQUENCE	RT	embryo. Its identity with the blood clotting factor Xa."	
FT	SEQUENCE	444	444	SEQUENCE	RL	FEBS Lett. 283:281-285(1991).	
FT	SEQUENCE	444	444	SEQUENCE	RN	[2]	
FT	SEQUENCE	444	444	SEQUENCE	RP	SEQUENCE OF 41-55 AND 241-261.	
FT	SEQUENCE	444	444	SEQUENCE	RC	TISSUE=Allantoic fluid;	

RA Negai Y.;  
 RT "An endoprotease homologous to the blood clotting factor X as a  
 RL EMBO J. 9:4189-4195(1990).  
 CC -|- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that  
 CC converts prothrombin to thrombin in the presence of factor Va,  
 CC calcium and phospholipid during blood clotting.  
 CC -|- FUNCTION: VAP CLEAVES THE FUSION PROTEINS OF SENDAI VIRUS, NDV,  
 CC AND INFLUENZA VIRUS A AT A SPECIFIC SINGLE ARGININE-CONTAINING  
 CC SITE, AND PLAYS A KEY ROLE IN THE VIRAL SPREADING IN THE ALLANTOIC  
 CC SAC.  
 CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then  
 CC Arg-|-Ile bonds in prothrombin to form thrombin.  
 CC -|- SUBUNIT: THE TWO CHAINS ARE FORMED FROM A SINGLE-CHAIN PRECURSOR  
 CC BY THE EXCISION OF TWO ARG RESIDUES AND ARE HELD TOGETHER BY 1 OR  
 CC MORE DISULFIDE BONDS.  
 CC -|- TISSUE SPECIFICITY: LIVER AND CHORIOALLANTOIC MEMBRANE.  
 CC -|- PTM: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME  
 CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND  
 CC CALCIUM.  
 CC -|- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE  
 CC INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).  
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -|- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or  
 CC send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: D00844; BAA00724.1; -  
 CC PIR: S15838; S15838.  
 CC HSP: P00742; IHCG.  
 CC MEROPS: S01.216; -  
 CC InterPro: IPR000152; Asx\_hydroxyl.  
 CC InterPro: IPR001314; Chymotrypsin.  
 CC InterPro: IPR000561; EGF-like.  
 CC InterPro: IPR000742; EGF-2.  
 CC InterPro: IPR011881; EGF-Ca.  
 CC InterPro: IPR001436; EGF-II.  
 CC InterPro: IPR002383; GLA\_blood.  
 CC InterPro: IPR001254; Ser\_protease\_Try.  
 CC InterPro: IPR000294; VitK\_dep\_GLA.  
 CC Pfam: PF00008; EGF; 2.  
 CC Pfam: PF00089; trypsin; 1.  
 CC Pfam: PF00594; gla; 1.  
 CC PRINTS: PR00722; CHYMOTRYPSIN.  
 CC PRINTS: PR00010; EGF\_BLOOD.  
 CC PRINTS: PR00001; GLABLOOD.  
 CC SMART: SM00179; EGF\_CA; 1.  
 CC SMART: SM00069; GLA; 1.  
 CC SMART: SM00020; Tryp\_Spc; 1.  
 CC PROSITE: PS00010; ASX\_HYDROXYL; 1.  
 CC PROSITE: PS00022; EGF\_1; 1.  
 CC PROSITE: PS01186; EGF\_2; 2.  
 CC PROSITE: PS01187; EGF\_CA; 1.  
 CC PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 CC PROSITE: PS00240; TRYPSIN\_DOM; 1.  
 CC PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE: PS00135; TRYPSIN\_SER; 1.  
 CC Glycoprotein; Hydrolase; Serine protease; Blood coagulation;  
 CC Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;  
 CC Signal; Zymogen; EGF-like domain; Repeat.  
 CC SIGNAL 1 20 OR 30, OR 31 (POTENTIAL).  
 CC PROPEP 1 20  
 CC CHAIN 21 40  
 CC CHAIN 41 180  
 CC CHAIN 186 475  
 CC PROPEP 186 241  
 CC CHAIN 242 475  
 CC CHAIN 86 122  
 CC DOMAIN

FT	DOMAIN	125	168	EGF-LIKE 2.
FT	DOMAIN	241	475	SERINE PROTEASE.
FT	MOD_RES	46	46	GAMMA-CARBOXYGLUTAMIC ACID
FT	MOD_RES	47	47	(BY SIMILARITY).
FT	MOD_RES	47	47	GAMMA-CARBOXYGLUTAMIC ACID
FT	MOD_RES	54	54	(BY SIMILARITY).
FT	MOD_RES	54	54	GAMMA-CARBOXYGLUTAMIC ACID
FT	MOD_RES	56	56	(BY SIMILARITY).
FT	MOD_RES	56	56	GAMMA-CARBOXYGLUTAMIC ACID
FT	MOD_RES	59	59	(BY SIMILARITY).
FT	MOD_RES	59	59	GAMMA-CARBOXYGLUTAMIC ACID
FT	MOD_RES	60	60	(BY SIMILARITY).
FT	MOD_RES	60	60	GAMMA-CARBOXYGLUTAMIC ACID
FT	MOD_RES	65	65	(BY SIMILARITY).
FT	MOD_RES	65	65	GAMMA-CARBOXYGLUTAMIC ACID
FT	MOD_RES	66	66	(BY SIMILARITY).
FT	MOD_RES	66	66	GAMMA-CARBOXYGLUTAMIC ACID
FT	MOD_RES	69	69	(BY SIMILARITY).
FT	MOD_RES	69	69	GAMMA-CARBOXYGLUTAMIC ACID
FT	MOD_RES	72	72	(BY SIMILARITY).
FT	MOD_RES	72	72	GAMMA-CARBOXYGLUTAMIC ACID
FT	MOD_RES	79	79	(BY SIMILARITY).
FT	MOD_RES	79	79	GAMMA-CARBOXYGLUTAMIC ACID
FT	MOD_RES	103	103	(BY SIMILARITY).
FT	MOD_RES	103	103	HYDROXYLATION (BY SIMILARITY).
FT	ACT_SITE	282	282	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	328	328	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	425	425	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID	90	101	BY SIMILARITY.
FT	DISULFID	95	110	BY SIMILARITY.
FT	DISULFID	112	121	BY SIMILARITY.
FT	DISULFID	129	140	BY SIMILARITY.
FT	DISULFID	136	152	BY SIMILARITY.
FT	DISULFID	154	167	BY SIMILARITY.
FT	DISULFID	175	348	INTERCHAIN (BY SIMILARITY).
FT	DISULFID	247	252	BY SIMILARITY.
FT	DISULFID	267	283	BY SIMILARITY.
FT	DISULFID	396	410	BY SIMILARITY.
FT	DISULFID	421	449	BY SIMILARITY.
FT	CARBOHYD	196	196	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	207	207	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	228	228	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	285	285	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SEQUENCE	475 AA;	53142 MW;	570BF84956C5C74D CRG64;

Alignment Scores:  
 Pred. No.: 1.4e-35  
 Score: 801.50  
 Percent Similarity: 54.77%  
 Best Local Similarity: 36.14%  
 Query Match: 34.30%  
 DB: 1  
 Gaps: 8

US-09-763-153-1 (1-1245) x FA10\_CHICK (1-475)

QY	1	CCCAACTCTCTTCGTGGAGGAGCTTCGTCCACAGCAGCTGGAGCGGAGTGCATAGAGGAG	60
Db	41	AlaAsnSerPheLeuGluGluMetLysGlnGlyAsnIleGluArgGluCysAsnGluGlu	60
QY	61	ATCTGTGACTTCGAGGAGCCCAAGGAATTTCCAAATGTGGATGACACTGGCCTTC	120
Db	61	ArgCysSerLysGluGluAlaArgGluAlaPheGluAspAsnGluLysThrGluGluPhe	80
QY	121	TGTTCCACAGCTCGACGGTGCAGCTGCTGTGCTTGGCTTGGACACCGTCGCGC	180
Db	81	TrpAsnIleTyrValAspGlyAspGlnCysSerSerAsnPro	94
QY	181	AGCCTGTGCTGGCGGACGACGCTGCATCGACGATCGGACGCTTACGTCGCGCTGC	240
Db	95	-----CysHisTyrGlyGlnCysLysAspGlyLeuGlySerTyrThrCysSerCys	112
QY	241	CCGACGCGCTGGAGGCGCGCTTCTGCCACGCGAGGTGAGCTTCTCAATTCGCTG	300



Db 113 LeuAspGlyTyrGlnGlyLysAsnCysGluPheValIlePro---LysTyrCysLysIle 131  
QY 301 GACACGCGCTGCACGCGATTACTGCTAGAGAGAGTGGCTGGCGGCGC----- 351  
Db 132 AsnAsnGlyAspCysGluGlnPheCysSerIleLysLysSerValGlnLysAspValVal 151  
QY 352 TGTAGCTGTGGCTGCTACAGCTGGGGAGACACCTCTCTGAGTGTACCCCGGAGTG 411  
Db 152 CysSerCysThrSerGlyTyrGluLeuAlaGluAspGlyLysGlnCysValSerLysVal 171  
QY 412 AAGTTCCTCTGTGGAGCGCTGGAAGCGGATGAGAGAGCGAGTCACTCAAAACGA 471  
Db 172 LysTyrProCysGlyLysValLeuMetLysArgIleLysArgSerValIleLeuProThr 191  
QY 472 GAC-----ACAGAGACCAA----- 486  
Db 192 AsnSerAsnThrAsnAlaThrSerAspGlnAspValProSerThrAsnGlySerIleLeu 211  
QY 486 ----- 486  
Db 212 GluGluValPheThrThrThrThrGluSerProThrProProArgAsnGlySerSer 231  
QY 487 -----GAAGACCAAGTAGATCCCGCTCATTCATGGGAAGATGACCGCGGGAGAC 540  
Db 232 IleThrAspProAsnValAspThrArgIleValGlyGlyAspGluCysArgProGlyGlu 251  
QY 541 AGCCCTGTGGCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
Db 252 CysProThrGlnAlaValLeuIleAsnGluLysGlyGluGluPheCysGlyGlyThrIle 271  
QY 601 ATCCACCCCTCTCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
Db 272 LeuAsnGluAspPheIleLeuThrAlaAlaHisCysIleAsnGlnSerLysGluIleLys 291  
QY 661 GTCAGCTGTGAGAGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
Db 292 ValValValGluValValAspArgGluLysGluGluHisSerGluThrHisThrAla 311  
QY 721 AAGGAGGTCTGCTGCCACCCCACTACAGCAAGAGCACCACCGACATGATGATGCTGCT 780  
Db 312 GluLysIlePheValHisSerLysTyrIleAlaGluThrTyrAspAsnAspIleAlaLeu 331  
QY 781 CTGACCTGCGCCGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 332 IleLysLeuLysGluProIleGlnPheSerGluThrValValProAlaCysLeuProGln 351  
QY 841 AGCGCCCTGTGAGAGCGCGAGCTC---AATCAGCGCCGCGCAGGAGACCTCTGCTGCTGCT 897  
Db 352 AlaAspPheAlaAsnGluValLeuMetAsnGln-----LysSerGlyMetValSerGly 369  
QY 898 TGGGCTACCCAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957  
Db 370 PheGlyArgGluPheGluAlaGlyArgLeuSerLysArg-----LeuLys 384  
QY 958 TTCATCAAGATCCCGTGTCCGCGACATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017  
Db 385 ValLeuGluValProTyrValAspArgSerThrCysLysGlnSerThrAsnPheAlaIle 404  
QY 1018 TCTGAGACATGCTGTGCTGGGCGATCTCTGGGAGCAGCGAGAGTGTCTGCTGCTGCTGCT 1077  
Db 405 ThrGluAsnMetPheCysAlaGlyTyrGluThrGluGlnLysAspAlaCysGlnGlyAsp 424  
QY 1078 AGTGGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137  
Db 425 SerGlyGlyProHisValThrArgTyrLysAspThrThrPheValThrGlyIleValSer 444  
QY 1138 TGGGCTGAGGCTGTGGGCTCTCTCAACTACGGCGTGTACCAAAAGTACAGCGCTAC 1197  
Db 445 TrpGlyGluGlyCysAlaArgLysGlyLysTyrGlyValTyrThrLysLeuSerArgPhe 464  
QY 1198 CTCAGCTGATCGATGGGCGACATCAGACACAG 1230  
Db 465 LeuArgTrpValArgThrValMetArgGlnLys 475

RESULT 12  
FA10\_HUMAN  
ID FA10\_HUMAN STANDARD; PRT; 488 AA.  
AC P00742; Q14340;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).  
GN F10.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91216473; PubMed=1902434;  
RA Messler T.L., Pittman D.D., Long G.L., Kaufman R.J., Church W.R.;  
RT "Cloning and expression in COS-1 cells of a full-length cDNA encoding  
RT human coagulation factor X";  
RL Gene 99:291-294(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87026600; PubMed=3768336;  
RA Leytus S.P., Foster D.C., Kurachi K., Davie E.W.;  
RT "Gene for human factor X: a blood coagulation factor whose gene  
RT organization is essentially identical with that of factor IX and  
RT protein C";  
RL Biochemistry 25:5098-5102(1986).  
RN [3]  
RP SEQUENCE OF 13-488 FROM N.A.  
RX MEDLINE=85216545; PubMed=2582420;  
RA Fung M.R., Hay C.W., McGillivray R.T.A.;  
RT "Characterization of an almost full-length cDNA coding for human  
RT blood coagulation factor X";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:3591-3595(1985).  
RN [4]  
RP SEQUENCE OF 19-488 FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=86221713; PubMed=3011603;  
RA Kaul R.K., Hildebrand B., Roberts S., Jagadeeswaran P.;  
RT "Isolation and characterization of human blood-coagulation factor X  
RT cDNA";  
RL Gene 41:311-314(1986).  
RN [5]  
RP SEQUENCE OF 41-179  
RX MEDLINE=83257207; PubMed=6871167;  
RA McMullen B.A., Fujikawa K., Kisiel W., Sasagawa T., Howald W.N.,  
RA Kwa E.Y., Weinstein B.;  
RT "Complete amino acid sequence of the light chain of human blood  
RT coagulation factor X: evidence for identification of residue 63 as  
RT beta-hydroxyaspartic acid";  
RL Biochemistry 22:2875-2884(1983).  
RN [6]  
RP SEQUENCE OF 115-488 FROM N.A., AND TISSUE SPECIFICITY.  
RC TISSUE=Liver;  
RX MEDLINE=84222026; PubMed=6587384;  
RA Leytus S.P., Chung D.W., Kisiel W., Kurachi K., Davie E.W.;  
RT "Characterization of a cDNA coding for human factor X";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:3699-3702(1984).  
RN [7]  
RP SEQUENCE OF 183-234, AND CARBOHYDRATE-LINKAGE SITES.  
RX MEDLINE=94062825; PubMed=8243461;  
RA Inoue K., Morita T.;  
RT "Identification of O-linked oligosaccharide chains in the activation  
RT peptides of blood coagulation factor X. The role of the carbohydrate  
RT moieties in the activation of factor X";  
RL Eur. J. Biochem. 218:153-163(1993).  
RN [8]  
RP SEQUENCE OF 1-23 FROM N.A.  
RX MEDLINE=90128299; PubMed=2612918;  
RA Jagadeeswaran P., Reddy S.V., Rao K.J., Hamsabhusan K., Lyman G.;  
RT "Cloning and characterization of the 5' end (exon 1) of the gene

RT encoding human factor X.";  
 RL Gene 84:517-519(1989).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 86-179 AND 235-278.  
 RX MEDLINE-93360277; PubMed-8355279;  
 RA Padmanabhan K., Padmanabhan K.P., Tulinsky A., Park C.H., Bode W.,  
 RA Huber R., Blankenship D.T., Cardin A.D., Kistiel W.,  
 RT "Structure of human Des(1-45) factor Xa at 2.2-A resolution.";  
 RL J. Mol. Biol. 232:947-966(1993).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 86-179 AND 235-278.  
 RX MEDLINE-98283982; PubMed-9618463;  
 RA Kamata K., Kawamoto H., Homma T., Iwama T., Kim S.H.;  
 RT "Structural basis for chemical inhibition of human blood coagulation  
 factor Xa";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:6630-6635(1998).  
 CC -|- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that  
 CC converts prothrombin to thrombin in the presence of factor Va,  
 CC calcium and phospholipid during blood clotting.  
 CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then  
 CC Arg-|-Ile bonds in prothrombin to form thrombin.  
 CC -|- SUBUNIT: THE TWO CHAINS ARE FORMED FROM A SINGLE-CHAIN PRECURSOR  
 CC BY THE EXCISION OF TWO ARG RESIDUES AND ARE HELD TOGETHER BY 1 OR  
 CC MORE DISULFIDE BONDS.  
 CC -|- TISSUE SPECIFICITY: Plasma; synthesized in the liver.  
 CC -|- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME  
 CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND  
 CC CALCIUM.  
 CC -|- PTM: N- AND O-GLYCOSYLATED.  
 CC -|- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE  
 CC INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).  
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -|- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; K03194; AAA52490.1; -;  
 DR EMBL; M57285; AAA52421.1; -;  
 DR EMBL; L29433; AAA52764.1; -;  
 DR EMBL; L00390; AAA52764.1; JOINED.  
 DR EMBL; L00391; AAA52764.1; JOINED.  
 DR EMBL; L00392; AAA52764.1; JOINED.  
 DR EMBL; L00393; AAA52764.1; JOINED.  
 DR EMBL; L00394; AAA52764.1; JOINED.  
 DR EMBL; L00395; AAA52764.1; JOINED.  
 DR EMBL; L00396; AAA52764.1; JOINED.  
 DR EMBL; M2613; AAA51984.1; -;  
 DR EMBL; K01886; AAA52486.1; -;  
 DR EMBL; M33297; AAA52636.1; -;  
 DR PIR; A00924; EXHU.  
 DR PIR; A25853; A25853.  
 DR PIR; A24478; A24478.  
 DR PDB; IHCG; 08-MAY-95.  
 DR PDB; IFAX; 29-OCT-97.  
 DR PDB; IFXY; 17-JUN-98.  
 DR PDB; IXKA; 23-MAR-99.  
 DR PDB; IXKB; 23-MAR-99.  
 DR MEROPS; S01.216; -;  
 DR GlycoSuiteDB; P00742; -;  
 DR Genew; HGNC:3528; F10.  
 DR MIM; 134530; -;  
 DR MIM; 227600; -;  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR000742; EGF\_2.  
 DR InterPro; IPR001881; EGF\_Ca.

DR InterPro; IPR002383; GLA\_blood.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR InterPro; IPR000294; VitK\_dep\_GLA.  
 DR Pfam; PF00008; EGF\_2.  
 DR Pfam; PF00089; trypsin; 1.  
 DR Pfam; PF00594; gla; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00001; GLABLOOD.  
 DR SMART; SM00179; EGF\_CA; 1.  
 DR SMART; SM00001; EGF\_like; 1.  
 DR SMART; SM00069; GLA; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 2.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 DR PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Glycoprotein; Hydrolase; Serine protease; Plasma; Blood coagulation;  
 KW gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;  
 KW Signal; zymogen; EGF-like domain; Repeat; 3D-structure.  
 FT SIGNAL 1 31  
 FT PROPEP 32 40  
 FT CHAIN 41 179  
 FT CHAIN 183 488  
 FT PROPEP 183 234  
 FT CHAIN 235 488  
 FT DOMAIN 86 122  
 FT DOMAIN 125 165  
 FT DOMAIN 235 488  
 FT MOD\_RES 46 46  
 FT MOD\_RES 47 47  
 FT MOD\_RES 54 54  
 FT MOD\_RES 56 56  
 FT MOD\_RES 59 59  
 FT MOD\_RES 60 60  
 FT MOD\_RES 65 65  
 FT MOD\_RES 66 66  
 FT MOD\_RES 69 69  
 FT MOD\_RES 72 72  
 FT MOD\_RES 79 79  
 FT MOD\_RES 103 103  
 FT CARBOHYD 199 199  
 FT CARBOHYD 211 211  
 FT CARBOHYD 221 221  
 FT CARBOHYD 231 231  
 FT ACT\_SITE 276 276  
 FT ACT\_SITE 322 322  
 FT ACT\_SITE 419 419  
 FT DISULFID 90 101  
 FT DISULFID 95 110  
 FT DISULFID 112 121  
 FT DISULFID 129 140  
 Alignment Scores:  
 Pred. No.: 14e-35 Length: 488  
 Score: 801.50 Matches: 159  
 Percent Similarity: 54.55% Conservative: 87  
 Best Local Similarity: 35.25% Mismatches: 150  
 Query Match: 34.30% Indels: 55  
 DB: 1 Gaps: 8

US-09-763-153-1 (1-1245) x FA10\_HUMAN (1-488)

QY 1 GCCAACTCTCTCTGAGGAGTCCGTCACAGCGGAGGAGTCATAGAGGAG 60  
 |||||||||||||||||||  
 Db 41 AlaAsnSerPheLeuGluGluMetLysGlyHisLeuGluArgLysCysMetGluGlu 60  
 |||||||||||||||||||  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGGAAATTTCCAAAAATGTGGATGCACACTGGCCTTC 120

[illegible]

RT (Xyl2-Glc) O-glycosidically linked to a serine residue in the first  
 RT epidermal growth factor-like domain of human factors VII and IX and  
 RT protein Z and bovine protein Z".  
 RL J. Biol. Chem. 264:20320-20325(1989).  
 RN [7]  
 RP STRUCTURE OF CARBOHYDRATE ON SER-112.  
 RX MEDLINE=91344709; PubMed=2129367;  
 RA Iwanaga S., Nishimura H., Kawabata S., Kisiel W., Hase S., Ikenaka T.;  
 RA "A new trisaccharide sugar chain linked to a serine residue in the  
 RT first EGF-like domain of clotting factors VII and IX and protein Z";  
 RL Adv. Exp. Med. Biol. 281:121-131(1990).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF FVIIA IN COMPLEX WITH TF.  
 RX MEDLINE=96175641; PubMed=8598903;  
 RA Banner D.W., D'Arcy A., Chene C., Winkler F.K., Guha A.,  
 RA Konigsberg W.H., Nemerson Y., Kirchhofer D.;  
 RA "The crystal structure of the complex of blood coagulation factor  
 RT VIIa with soluble tissue factor";  
 RL Nature 380:41-46(1996).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF FVIIA IN COMPLEX WITH TF.  
 RX MEDLINE=99126538; PubMed=9925787;  
 RA Zhang E., St Charles R., Tulinsky A.;  
 RA "Structure of extracellular tissue factor complexed with factor VIIa  
 RT inhibited with a BPTI mutant";  
 RL J. Mol. Biol. 285:2089-2104(1999).  
 RN [10]  
 RP STRUCTURE BY NMR OF 105-145.  
 RX MEDLINE=98367502; PubMed=9692950;  
 RA Muranyi A., Finn B.E., Gippert G.P., Forsen S., Stenflo J.,  
 RA Drakenberg T.;  
 RA "Solution structure of the N-terminal EGF-like domain from human  
 RT factor VII";  
 RL Biochemistry 37:10605-10615(1998).  
 RN [11]  
 RP VARIANT GLN-364.  
 RX MEDLINE=91300046; PubMed=2070047;  
 RA O'Brien D.P., Gale K.M., Anderson J.S., McVey J.H., Miller G.J.,  
 RA Meade T.W., Tuddenham E.G.D.;  
 RA "Purification and characterization of factor VII 304-Gln: a variant  
 RT molecule with reduced activity isolated from a clinically unaffected  
 RL male";  
 RL Blood 78:132-140(1991).  
 RN [12]  
 RP VARIANTS GLN-364 AND PHE-370.  
 RX MEDLINE=92340074; PubMed=1634227;  
 RA Marchetti G., Patraccchini P., Gemmati D., Derosa V., Pinotti M.,  
 RA Rodorigo G., Casonato A., Girolami A., Bernardi F.;  
 RA "Detection of two missense mutations and characterization of a repeat  
 RT polymorphism in the factor VII gene (F7)".  
 RL Hum. Genet. 89:497-502(1992).  
 RN [13]  
 RP VARIANT TYR-238.  
 RX MEDLINE=93372811; PubMed=8364544;  
 RA Marchetti G., Ferrati M., Patraccchini P., Redaelli P., Bernardi F.;  
 RA "A missense mutation (178Cys-->Tyr) and two neutral dimorphisms  
 RT (115His and 333Ser) in the human coagulation factor VII gene";  
 RL Hum. Mol. Genet. 2:1055-1056(1993).  
 RN [14]  
 RP VARIANTS.  
 RX MEDLINE=94061028; PubMed=8242057;  
 RA Takamiya O., Kemball-Cook G., Marin D.M.A., Cooper D.N.,  
 RA von Felten A., Meili E., Hahn I., Prangnell D.R., Lumley H.,  
 RA Tuddenham E.G.D., McVey J.H.;  
 RA "Detection of missense mutations by single-strand conformational  
 RT polymorphism (SSCP) analysis in five dysfunctional variants of  
 RL coagulation factor VII";  
 RL Hum. Mol. Genet. 2:1355-1359(1993).  
 RN [15]  
 RP VARIANTS CHARLOTTE GLN-139 AND GLN-212.  
 RX MEDLINE=94264305; PubMed=8204879;  
 RA Chaing S., Clarke B., Sridhara S., Chu K., Friedman P., Vandusen W.,  
 RA Roberts H.R., Blajchman M., Monroe D.M., High K.A.;

RT "Severe factor VII deficiency caused by mutations abolishing the  
 RT cleavage site for activation and altering binding to tissue factor";  
 RL Blood 83:3524-3535(1994).  
 RN [16]  
 RP VARIANT VAL-354.  
 RX MEDLINE=95072589; PubMed=7981691;  
 RA Bernardi F., Castaman G., Redaelli R., Pinotti M., Lunghi B.,  
 RA Rodeghiero F., Marchetti G.;  
 RA "Topologically equivalent mutations causing dysfunctional coagulation  
 RT factors VII (294Ala-->Val) and X (334Ser-->Pro)";  
 RL Hum. Mol. Genet. 3:1175-1177(1994).  
 RN [17]  
 RP VARIANT MIE HIS-307.  
 RX MEDLINE=95064662; PubMed=7974346;  
 RA Ohiwa M., Hayashi T., Wada H., Minamikawa K., Shirakawa S.,  
 RA Suzuki K.;  
 RA "Factor VII MIE: homozygous asymptomatic type I deficiency caused by  
 RT an amino acid substitution of His (CAC) for Arg(247) (CGC) in the  
 RL catalytic domain";  
 RL Thromb. Haemost. 71:773-777(1994).  
 RN [18]  
 RP VARIANT MET-419.  
 RX MEDLINE=96247510; PubMed=8652821;  
 RA Arbini A.A., Mannucci P.M., Bauer K.A.;  
 RA "A Thr359Met mutation in factor VII of a patient with a hereditary  
 RT deficiency causes defective secretion of the molecule";  
 RL Blood 87:5085-5094(1996).  
 RN [19]  
 RP VARIANTS W-283; K-325; V-358; Q-364; E-402 AND Q-413.  
 RX MEDLINE=97001216; PubMed=8844208;  
 RA Bernardi F., Castaman G., Pinotti M., Ferraresi P., di Iasio M.G.,  
 RA Lunghi B., Rodeghiero F., Marchetti G.;  
 RA "Mutation pattern in clinically asymptomatic coagulation factor VII  
 RT deficiency";  
 RL Hum. Mutat. 8:108-115(1996).  
 RN [20]  
 RP VARIANT VAL-304.  
 RX MEDLINE=97037613; PubMed=8883260;  
 RA Tanary H., Fromovich Y., Shalmon L., Reich Z., Dym O., Lanir N.,  
 RA Brenner B., Paz M., Luder A.S., Blau O., Korostishevsky M.,  
 RA Zaizov R., Seligsohn U.;  
 RA "Ala244Val is a common, probably ancient mutation causing factor VII  
 RT deficiency in Moroccan and Iranian Jews";  
 RL Thromb. Haemost. 76:283-291(1996).  
 RN [21]  
 RP VARIANTS MALTA THR-194 AND VAL-304.  
 RX MEDLINE=98112461; PubMed=9452082;  
 RA Alshinawi C., Scerri C., Gaidies R., Aquilina A., Felice A.E.;  
 RA "Two new missense mutations (P134T and A244V) in the coagulation  
 RT factor VII gene";  
 RL Hum. Mutat. Suppl. 1:S189-S191(1998).  
 RN [22]  
 RP FUNCTION: CIRCULATES IN THE BLOOD IN A ZMOGEN FORM. FACTOR VII IS  
 CC CONVERTED TO FACTOR VIIA BY FACTOR XA, FACTOR XIIA, FACTOR IXA, OR  
 CC THROMBIN BY MINOR PROTEOLYSIS. IN THE PRESENCE OF TISSUE FACTOR  
 CC AND CALCIUM IONS, FACTOR VIIA THEN CONVERTS FACTOR X TO FACTOR XA  
 CC BY LIMITED PROTEOLYSIS. FACTOR VIIA WILL ALSO CONVERT FACTOR IX TO  
 CC FACTOR IXA IN THE PRESENCE OF TISSUE FACTOR AND CALCIUM.  
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes one Arg-Ile bond in factor X to  
 CC form factor Xa.  
 CC -1- SUBUNIT: HETERODIMER OF A LIGHT CHAIN AND A HEAVY CHAIN LINKED  
 CC BY A DISULFIDE BOND.  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; A (shown here) and B; are  
 CC produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: PLASMA.  
 CC -1- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME  
 CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND  
 CC CALCIUM.  
 CC -1- DISEASE: DEFECTS IN F7 CAN CAUSE COAGULOPATHY.  
 CC -1- PHARMACEUTICAL: Available under the names Niasase or Novoseven  
 CC (Novo Nordisk). Used for the treatment of bleeding episodes in

Alignment Scores:

Pred. No.:

1.33e-34

Length:

466

Score:	783.00	Matches:	164
Percent Similarity:	56.74%	Conservative:	76
Best Local Similarity:	38.77%	Mismatches:	147
Query Match:	33.50%	Indels:	36
DB:	1	Gaps:	10

  

US-09-763-153-1 (1-1245) x FA7_HUMAN (1-466)	
QY 1 GCCAACCTCTCTGGAGGAGCTCCATCAGCAGCGCTGGAGCGGAGTGCATAGGAG 60	DB 355 LeuGluMetValLeuAsnValProArgLeuMetThrGlnAspCysLeuGlnGlnSer 374
DB 61 AlaAsnAlaPheLeuGluGluLeuArgProGlySerLeuGluArgGlnCysLeuGlu 80	QY 1006 -----AGCAACATGGTGTGAGAACATGCTGTGGCGGCGCATCTCTCGG 1050
QY 61 ATCTGTGACATCGAGGAGGCAAGAAATTTCCAAATTTGGATGACACACTGCCCTTC 120	DB 375 ArgLysValGlyAspSerProAsnIleThrGluTyrMetPheCysAlaGlyTyrSerAsp 394
QY 81 GlnCysSerPheGluGluAlaArgGluIlePheLysAspAlaGluArgThrLysLeuPhe 100	QY 1051 GACCGCAGGATCCGTCGAGCGACAGTGGGGGCGCCATCGTCCCTCCACACGCG 1110
DB 121 TGGTCCAAAGCAGCTGCGAGTGCACAGTGTGGTCTGCTGCTGCGGACCGCTGCGC 180	DB 395 GlySerLysAspSerCysGlyAspSerGlyGlyProHisAlaThrHisTyrArgGly 414
QY 101 TrpIleSerTyrSerAspGlyAspGlnCys -----AlaSer 112	QY 1111 ACCTGGTTCCTGGTGGCGCTGTGAGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 1170
QY 181 AGCTGTGCTCGCGCAGCGCAGCTGCATCGAGCGCATCGCGCAGCTTCAGCTCGACTGC 240	DB 415 ThrTrpTyrLeuThrGlyIleValSerTrpGlyGlnGlyCysAlaThrValGlyHisPhe 434
DB 113 SerProCysGlnAsnGlyGlySerCysLysAspGlnLeuGlnSerTyrIleCysPheCys 132	QY 1171 GCGTTTACCAACAGCAGCGCTACCTCGACTGCATCGGCGGACATGAGACAGACAA 1230
QY 241 CGCAGCGCTGGAGGCGCTTCTGCCAG ---CGCAGGTGAGCTTCCTCAATTGCTCG 297	DB 435 GlyValTyrThrArgValSerGlnTyrIleGluTrpLeuGlnLysLeuMetArgSerGlu 454
DB 133 LeuProAlaPheGluGlyArgAsnCysGluThrHisLysAspAspGlnLeuIleCysVal 152	QY 1231 GAAGCCCC 1239
QY 298 CTGGACACGGCGCTGACGATCTACTGCTAGAGAGGTGGCTGGCGCGC ---TGT 354	DB 455 ProArgPro 457
DB 153 AsnGluAsnGlyGlyCysGluGlnTyrSerHisThrGlyThrLysArgSerCys 172	RESULT 14
QY 355 AGCTGTGCGCTGTACAGCTGGGCGACGACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 414	FA7_BOVIN
DB 173 ArgCysHisGluGlyTyrSerLeuLeuAlaAspGlyValSerCysThrProThrValGlu 192	ID FA7_BOVIN STANDARD; PRT; 407 AA.
QY 415 TTCCCTGTGTGGAGCGCTTGGAGCGGTATGGAGAGAGCGCAGCTCAGCTGAAACGAGAC 474	AC P22457;
DB 193 TyrProCysGlyLys ---IleProIleLeuGluLysArgAsnAla ----- 206	DT 01-AUG-1991 (Rel. 19, Created)
QY 475 ACAGAGACCAAGACCAAGTAGATCGCGCTCATGTGATGGAAGATGACACCGCGG 534	DT 01-AUG-1991 (Rel. 19, Last sequence update)
DB 207 -----SerLysProGlnGlyArgIleValGlyLysValCysProLys 221	DT 15-JUN-2002 (Rel. 41, Last annotation update)
QY 535 GGAGACAGCCCTGGAGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594	DE Coagulation factor VII (EC 3.4.21.21) (Serum prothrombin conversion accelerator).
DB 222 GlyLysCysProIleGlnValLeuLeuValAsnGlyAlaGlnLeu ---CysGlyGly 240	GN F7.
QY 595 GTGCTATCCACCCCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 651	OS Bos taurus (Bovine).
DB 241 ThrLeuIleAsnThrIleTrpValValSerAlaAlaHisCysPheAspLysLysAsn 260	OC Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
QY 652 -----AAGTCTCTGTGAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 705	OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
DB 261 TrpArgAsnLeuIleAlaValLeuGlyHisAspLeuSerGluHisAspGlyAspGlu 280	OC Bovidae; Bovinae; Bos.
QY 706 CTGGACCTGGACATCAAGAGGTCTTGGTCCACCCCACTACAGAGACACCCAGC 765	OX NCBI_TaxID=9913;
DB 281 GlnSerArgArgValAlaGlnValIleProSerThrTyrValProGlyThrThrAsn 300	RN [1]
QY 766 ATGATCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 825	RP SEQUENCE.
DB 301 HisAspIleAlaLeuLeuArgLeuHisGlnProValValLeuThrAspHisValPro 320	RX MEDLINE=89008362; PubMed=3049594;
QY 826 ATCTGCTCCGCGAGCGGCTTGCAGAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 885	RA Takeya H., Kawabata S., Nakagawa K., Yamamichi Y., Miyata T.,
DB 321 LeuCysLeuProGluArgThrPheSerGluArgThrLeuAlaPheVal ---ArgPheSer 339	RA Iwanaga S.,
QY 886 CTGCTGACGGGTGGGTGCTACACAGCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 945	RA Miyata T., Iwanaga S., Takao T., Shimomishi Y., Ikenaka T.;
DB 340 LeuValSerGlyTrpGlnLeuLeuAspArgGlyAlaThrAla ----- 354	RA "A new trisaccharide sugar chain linked to a serine residue in bovine blood coagulation factors VII and IX.";
QY 946 TTGCTCTCACTTCATCAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1005	RL J. Biol. Chem. 263:14868-14877(1988).

CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND  
 CC CALCIUM.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC PIR: A31979; A31979.  
 DR HSP: P08709; Ibf9.  
 DR MEROPS: S01.215; -.  
 DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF-2.  
 DR InterPro: IPR001881; EGF-Ca.  
 DR InterPro: IPR001438; EGF-II.  
 DR InterPro: IPR002383; GLA\_blood.  
 DR InterPro: IPR001254; Ser\_protease\_Try.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF00089; trypsin; 1.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PRINTS: PR00010; EGF\_blood.  
 DR PRINTS: PR00001; GLABLOOD.  
 DR SMART: SM00179; EGF\_CA; 1.  
 DR SMART: SM00001; EGF\_like; 1.  
 DR SMART: SM00069; GLA; 1.  
 DR SMART: SM00020; Tryp\_Spc; 1.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 2.  
 DR PROSITE: PS01187; EGF\_CA; 1.  
 DR PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase: Serine protease; Blood coagulation; Zymogen; Glycoprotein;  
 KW Liver; Plasma; Vitamin K; Calcium-binding; Gamma-carboxyglutamic acid;  
 KW EGF-like domain; Repeat.  
 FT CHAIN 1 152  
 FT DOMAIN 153 407  
 FT CHAIN 407 475  
 FT DOMAIN 407 475  
 FT CHAIN 475 534  
 FT DOMAIN 475 534  
 FT CHAIN 534 594  
 FT DOMAIN 534 594  
 FT CHAIN 594 651  
 FT DOMAIN 594 651  
 FT CHAIN 651 705  
 FT DOMAIN 651 705  
 FT CHAIN 705 765  
 FT DOMAIN 705 765  
 FT CHAIN 765 825  
 FT DOMAIN 765 825  
 FT CHAIN 825 885  
 FT DOMAIN 825 885  
 FT CHAIN 885 929  
 FT DOMAIN 885 929  
 FT CHAIN 929 985  
 FT DOMAIN 929 985  
 FT CHAIN 985 1045  
 FT DOMAIN 985 1045  
 FT CHAIN 1045 1105  
 FT DOMAIN 1045 1105  
 FT CHAIN 1105 1165  
 FT DOMAIN 1105 1165  
 FT CHAIN 1165 1225  
 FT DOMAIN 1165 1225  
 FT CHAIN 1225 1285  
 FT DOMAIN 1225 1285  
 FT CHAIN 1285 1345  
 FT DOMAIN 1285 1345  
 FT CHAIN 1345 1405  
 FT DOMAIN 1345 1405  
 FT CHAIN 1405 1465  
 FT DOMAIN 1405 1465  
 FT CHAIN 1465 1525  
 FT DOMAIN 1465 1525  
 FT CHAIN 1525 1585  
 FT DOMAIN 1525 1585  
 FT CHAIN 1585 1645  
 FT DOMAIN 1585 1645  
 FT CHAIN 1645 1705  
 FT DOMAIN 1645 1705  
 FT CHAIN 1705 1765  
 FT DOMAIN 1705 1765  
 FT CHAIN 1765 1825  
 FT DOMAIN 1765 1825  
 FT CHAIN 1825 1885  
 FT DOMAIN 1825 1885  
 FT CHAIN 1885 1945  
 FT DOMAIN 1885 1945  
 FT CHAIN 1945 2005  
 FT DOMAIN 1945 2005  
 FT CHAIN 2005 2065  
 FT DOMAIN 2005 2065  
 FT CHAIN 2065 2125  
 FT DOMAIN 2065 2125  
 FT CHAIN 2125 2185  
 FT DOMAIN 2125 2185  
 FT CHAIN 2185 2245  
 FT DOMAIN 2185 2245  
 FT CHAIN 2245 2305  
 FT DOMAIN 2245 2305  
 FT CHAIN 2305 2365  
 FT DOMAIN 2305 2365  
 FT CHAIN 2365 2425  
 FT DOMAIN 2365 2425  
 FT CHAIN 2425 2485  
 FT DOMAIN 2425 2485  
 FT CHAIN 2485 2545  
 FT DOMAIN 2485 2545  
 FT CHAIN 2545 2605  
 FT DOMAIN 2545 2605  
 FT CHAIN 2605 2665  
 FT DOMAIN 2605 2665  
 FT CHAIN 2665 2725  
 FT DOMAIN 2665 2725  
 FT CHAIN 2725 2785  
 FT DOMAIN 2725 2785  
 FT CHAIN 2785 2845  
 FT DOMAIN 2785 2845  
 FT CHAIN 2845 2905  
 FT DOMAIN 2845 2905  
 FT CHAIN 2905 2965  
 FT DOMAIN 2905 2965  
 FT CHAIN 2965 3025  
 FT DOMAIN 2965 3025  
 FT CHAIN 3025 3085  
 FT DOMAIN 3025 3085  
 FT CHAIN 3085 3145  
 FT DOMAIN 3085 3145  
 FT CHAIN 3145 3205  
 FT DOMAIN 3145 3205  
 FT CHAIN 3205 3265  
 FT DOMAIN 3205 3265  
 FT CHAIN 3265 3325  
 FT DOMAIN 3265 3325  
 FT CHAIN 3325 3385  
 FT DOMAIN 3325 3385  
 FT CHAIN 3385 3445  
 FT DOMAIN 3385 3445  
 FT CHAIN 3445 3505  
 FT DOMAIN 3445 3505  
 FT CHAIN 3505 3565  
 FT DOMAIN 3505 3565  
 FT CHAIN 3565 3625  
 FT DOMAIN 3565 3625  
 FT CHAIN 3625 3685  
 FT DOMAIN 3625 3685  
 FT CHAIN 3685 3745  
 FT DOMAIN 3685 3745  
 FT CHAIN 3745 3805  
 FT DOMAIN 3745 3805  
 FT CHAIN 3805 3865  
 FT DOMAIN 3805 3865  
 FT CHAIN 3865 3925  
 FT DOMAIN 3865 3925  
 FT CHAIN 3925 3985  
 FT DOMAIN 3925 3985  
 FT CHAIN 3985 4045  
 FT DOMAIN 3985 4045  
 FT CHAIN 4045 4105  
 FT DOMAIN 4045 4105  
 FT CHAIN 4105 4165  
 FT DOMAIN 4105 4165  
 FT CHAIN 4165 4225  
 FT DOMAIN 4165 4225  
 FT CHAIN 4225 4285  
 FT DOMAIN 4225 4285  
 FT CHAIN 4285 4345  
 FT DOMAIN 4285 4345  
 FT CHAIN 4345 4405  
 FT DOMAIN 4345 4405  
 FT CHAIN 4405 4465  
 FT DOMAIN 4405 4465  
 FT CHAIN 4465 4525  
 FT DOMAIN 4465 4525  
 FT CHAIN 4525 4585  
 FT DOMAIN 4525 4585  
 FT CHAIN 4585 4645  
 FT DOMAIN 4585 4645  
 FT CHAIN 4645 4705  
 FT DOMAIN 4645 4705  
 FT CHAIN 4705 4765  
 FT DOMAIN 4705 4765  
 FT CHAIN 4765 4825  
 FT DOMAIN 4765 4825  
 FT CHAIN 4825 4885  
 FT DOMAIN 4825 4885  
 FT CHAIN 4885 4945  
 FT DOMAIN 4885 4945  
 FT CHAIN 4945 5005  
 FT DOMAIN 4945 5005  
 FT CHAIN 5005 5065  
 FT DOMAIN 5005 5065  
 FT CHAIN 5065 5125  
 FT DOMAIN 5065 5125  
 FT CHAIN 5125 5185  
 FT DOMAIN 5125 5185  
 FT CHAIN 5185 5245  
 FT DOMAIN 5185 5245  
 FT CHAIN 5245 5305  
 FT DOMAIN 5245 5305  
 FT CHAIN 5305 5365  
 FT DOMAIN 5305 5365  
 FT CHAIN 5365 5425  
 FT DOMAIN 5365 5425  
 FT CHAIN 5425 5485  
 FT DOMAIN 5425 5485  
 FT CHAIN 5485 5545  
 FT DOMAIN 5485 5545  
 FT CHAIN 5545 5605  
 FT DOMAIN 5545 5605  
 FT CHAIN 5605 5665  
 FT DOMAIN 5605 5665  
 FT CHAIN 5665 5725  
 FT DOMAIN 5665 5725  
 FT CHAIN 5725 5785  
 FT DOMAIN 5725 5785  
 FT CHAIN 5785 5845  
 FT DOMAIN 5785 5845  
 FT CHAIN 5845 5905  
 FT DOMAIN 5845 5905  
 FT CHAIN 5905 5965  
 FT DOMAIN 5905 5965  
 FT CHAIN 5965 6025  
 FT DOMAIN 5965 6025  
 FT CHAIN 6025 6085  
 FT DOMAIN 6025 6085  
 FT CHAIN 6085 6145  
 FT DOMAIN 6085 6145  
 FT CHAIN 6145 6205  
 FT DOMAIN 6145 6205  
 FT CHAIN 6205 6265  
 FT DOMAIN 6205 6265  
 FT CHAIN 6265 6325  
 FT DOMAIN 6265 6325  
 FT CHAIN 6325 6385  
 FT DOMAIN 6325 6385  
 FT CHAIN 6385 6445  
 FT DOMAIN 6385 6445  
 FT CHAIN 6445 6505  
 FT DOMAIN 6445 6505  
 FT CHAIN 6505 6565  
 FT DOMAIN 6505 6565  
 FT CHAIN 6565 6625  
 FT DOMAIN 6565 6625  
 FT CHAIN 6625 6685  
 FT DOMAIN 6625 6685  
 FT CHAIN 6685 6745  
 FT DOMAIN 6685 6745  
 FT CHAIN 6745 6805  
 FT DOMAIN 6745 6805  
 FT CHAIN 6805 6865  
 FT DOMAIN 6805 6865  
 FT CHAIN 6865 6925  
 FT DOMAIN 6865 6925  
 FT CHAIN 6925 6985  
 FT DOMAIN 6925 6985  
 FT CHAIN 6985 7045  
 FT DOMAIN 6985 7045  
 FT CHAIN 7045 7105  
 FT DOMAIN 7045 7105  
 FT CHAIN 7105 7165  
 FT DOMAIN 7105 7165  
 FT CHAIN 7165 7225  
 FT DOMAIN 7165 7225  
 FT CHAIN 7225 7285  
 FT DOMAIN 7225 7285  
 FT CHAIN 7285 7345  
 FT DOMAIN 7285 7345  
 FT CHAIN 7345 7405  
 FT DOMAIN 7345 7405  
 FT CHAIN 7405 7465  
 FT DOMAIN 7405 7465  
 FT CHAIN 7465 7525  
 FT DOMAIN 7465 7525  
 FT CHAIN 7525 7585  
 FT DOMAIN 7525 7585  
 FT CHAIN 7585 7645  
 FT DOMAIN 7585 7645  
 FT CHAIN 7645 7705  
 FT DOMAIN 7645 7705  
 FT CHAIN 7705 7765  
 FT DOMAIN 7705 7765  
 FT CHAIN 7765 7825  
 FT DOMAIN 7765 7825  
 FT CHAIN 7825 7885  
 FT DOMAIN 7825 7885  
 FT CHAIN 7885 7945  
 FT DOMAIN 7885 7945  
 FT CHAIN 7945 8005  
 FT DOMAIN 7945 8005  
 FT CHAIN 8005 8065  
 FT DOMAIN 8005 8065  
 FT CHAIN 8065 8125  
 FT DOMAIN 8065 8125  
 FT CHAIN 8125 8185  
 FT DOMAIN 8125 8185  
 FT CHAIN 8185 8245  
 FT DOMAIN 8185 8245  
 FT CHAIN 8245 8305  
 FT DOMAIN 8245 8305  
 FT CHAIN 8305 8365  
 FT DOMAIN 8305 8365  
 FT CHAIN 8365 8425  
 FT DOMAIN 8365 8425  
 FT CHAIN 8425 8485  
 FT DOMAIN 8425 8485  
 FT CHAIN 8485 8545  
 FT DOMAIN 8485 8545  
 FT CHAIN 8545 8605  
 FT DOMAIN 8545 8605  
 FT CHAIN 8605 8665  
 FT DOMAIN 8605 8665  
 FT CHAIN 8665 8725  
 FT DOMAIN 8665 8725  
 FT CHAIN 8725 8785  
 FT DOMAIN 8725 8785  
 FT CHAIN 8785 8845  
 FT DOMAIN 8785 8845  
 FT CHAIN 8845 8905  
 FT DOMAIN 8845 8905  
 FT CHAIN 8905 8965  
 FT DOMAIN 8905 8965  
 FT CHAIN 8965 9025  
 FT DOMAIN 8965 9025  
 FT CHAIN 9025 9085  
 FT DOMAIN 9025 9085  
 FT CHAIN 9085 9145  
 FT DOMAIN 9085 9145  
 FT CHAIN 9145 9205  
 FT DOMAIN 9145 9205  
 FT CHAIN 9205 9265  
 FT DOMAIN 9205 9265  
 FT CHAIN 9265 9325  
 FT DOMAIN 9265 9325  
 FT CHAIN 9325 9385  
 FT DOMAIN 9325 9385  
 FT CHAIN 9385 9445  
 FT DOMAIN 9385 9445  
 FT CHAIN 9445 9505  
 FT DOMAIN 9445 9505  
 FT CHAIN 9505 9565  
 FT DOMAIN 9505 9565  
 FT CHAIN 9565 9625  
 FT DOMAIN 9565 9625  
 FT CHAIN 9625 9685  
 FT DOMAIN 9625 9685  
 FT CHAIN 9685 9745  
 FT DOMAIN 9685 9745  
 FT CHAIN 9745 9805  
 FT DOMAIN 9745 9805  
 FT CHAIN 9805 9865  
 FT DOMAIN 9805 9865  
 FT CHAIN 9865 9925  
 FT DOMAIN 9865 9925  
 FT CHAIN 9925 9985  
 FT DOMAIN 9925 9985  
 FT CHAIN 9985 10045  
 FT DOMAIN 9985 10045  
 FT CHAIN 10045 10105  
 FT DOMAIN 10045 10105  
 FT CHAIN 10105 10165  
 FT DOMAIN 10105 10165  
 FT CHAIN 10165 10225  
 FT DOMAIN 10165 10225  
 FT CHAIN 10225 10285  
 FT DOMAIN 10225 10285  
 FT CHAIN 10285 10345  
 FT DOMAIN 10285 10345  
 FT CHAIN 10345 10405  
 FT DOMAIN 10345 10405  
 FT CHAIN 10405 10465  
 FT DOMAIN 10405 10465  
 FT CHAIN 10465 10525  
 FT DOMAIN 10465 10525  
 FT CHAIN 10525 10585  
 FT DOMAIN 10525 10585  
 FT CHAIN 10585 10645  
 FT DOMAIN 10585 10645  
 FT CHAIN 10645 10705  
 FT DOMAIN 10645 10705  
 FT CHAIN 10705 10765  
 FT DOMAIN 10705 10765  
 FT CHAIN 10765 10825  
 FT DOMAIN 10765 10825  
 FT CHAIN 10825 10885  
 FT DOMAIN 10825 10885  
 FT CHAIN 10885 10945  
 FT DOMAIN 10885 10945  
 FT CHAIN 10945 11005  
 FT DOMAIN 10945 11005  
 FT CHAIN 11005 11065  
 FT DOMAIN 11005 11065  
 FT CHAIN 11065 11125  
 FT DOMAIN 11065 11125  
 FT CHAIN 11125 11185  
 FT DOMAIN 11125 11185  
 FT CHAIN 11185 11245  
 FT DOMAIN 11185 11245  
 FT CHAIN 11245 11305  
 FT DOMAIN 11245 11305  
 FT CHAIN 11305 11365  
 FT DOMAIN 11305 11365  
 FT CHAIN 11365 11425  
 FT DOMAIN 11365 11425  
 FT CHAIN 11425 11485  
 FT DOMAIN 11425 11485  
 FT CHAIN 11485 11545  
 FT DOMAIN 11485 11545  
 FT CHAIN 11545 11605  
 FT DOMAIN 11545 11605  
 FT CHAIN 11605 11665  
 FT DOMAIN 11605 11665  
 FT CHAIN 11665 11725  
 FT DOMAIN 11665 11725  
 FT CHAIN 11725 11785  
 FT DOMAIN 11725 11785  
 FT CHAIN 11785 11845  
 FT DOMAIN 11785 11845  
 FT CHAIN 11845 11905  
 FT DOMAIN 11845 11905  
 FT CHAIN 11905 11965  
 FT DOMAIN 11905 11965  
 FT CHAIN 11965 12025  
 FT DOMAIN 11965 12025  
 FT CHAIN 12025 12085  
 FT DOMAIN 12025 12085  
 FT CHAIN 12085 12145  
 FT DOMAIN 12085 12145  
 FT CHAIN 12145 12205  
 FT DOMAIN 12145 12205  
 FT CHAIN 12205 12265  
 FT DOMAIN 12205 12265  
 FT CHAIN 12265 12325  
 FT DOMAIN 12265 12325  
 FT CHAIN 12325 12385  
 FT DOMAIN 12325 12385  
 FT CHAIN 12385 12445  
 FT DOMAIN 12385 12445  
 FT CHAIN 12445 12505  
 FT DOMAIN 12445 12505  
 FT CHAIN 12505 12565  
 FT DOMAIN 12505 12565  
 FT CHAIN 12565 12625  
 FT DOMAIN 12565 12625  
 FT CHAIN 12625 12685  
 FT DOMAIN 12625 12685  
 FT CHAIN 12685 12745  
 FT DOMAIN 12685 12745  
 FT CHAIN 12745 12805  
 FT DOMAIN 12745 12805  
 FT CHAIN 12805 12865  
 FT DOMAIN 12805 12865  
 FT CHAIN 12865 12925  
 FT DOMAIN 12865 12925  
 FT CHAIN 12925 12985  
 FT DOMAIN 12925 12985  
 FT CHAIN 12985 13045  
 FT DOMAIN 12985 13045  
 FT CHAIN 13045 13105  
 FT DOMAIN 13045 13105  
 FT CHAIN 13105 13165  
 FT DOMAIN 13105 13165  
 FT CHAIN 13165 13225  
 FT DOMAIN 13165 13225  
 FT CHAIN 13225 13285  
 FT DOMAIN 13225 13285  
 FT CHAIN 13285 13345  
 FT DOMAIN 13285 13345  
 FT CHAIN 13345 13405  
 FT DOMAIN 13345 13405  
 FT CHAIN 13405 13465  
 FT DOMAIN 13405 13465  
 FT CHAIN 13465 13525  
 FT DOMAIN 13465 13525  
 FT CHAIN 13525 13585  
 FT DOMAIN 13525 13585  
 FT CHAIN 13585 13645  
 FT DOMAIN 13585 13645  
 FT CHAIN 13645 13705  
 FT DOMAIN 13645 13705  
 FT CHAIN 13705 13765  
 FT DOMAIN 13705 13765  
 FT CHAIN 13765 13825  
 FT DOMAIN 13765 13825  
 FT CHAIN 13825 13885  
 FT DOMAIN 13825 13885  
 FT CHAIN 13885 13945  
 FT DOMAIN 13885 13945  
 FT CHAIN 13945 14005  
 FT DOMAIN 13945 14005  
 FT CHAIN 14005 14065  
 FT DOMAIN 14005 14065  
 FT CHAIN 14065 14125  
 FT DOMAIN 14065 14125  
 FT CHAIN 14125 14185  
 FT DOMAIN 14125 14185  
 FT CHAIN 14185 14245  
 FT DOMAIN 14185 14245  
 FT CHAIN 14245 14305  
 FT DOMAIN 14245 14305  
 FT CHAIN 14305 14365  
 FT DOMAIN 14305 14365  
 FT CHAIN 14365 14425  
 FT DOMAIN 14365 14425  
 FT CHAIN 14425 14485  
 FT DOMAIN 14425 14485  
 FT CHAIN 14485 14545  
 FT DOMAIN 14485 14545  
 FT CHAIN 14545 14605  
 FT DOMAIN 14545 14605  
 FT CHAIN 14605 14665  
 FT DOMAIN 14605 14665  
 FT CHAIN 14665 14725  
 FT DOMAIN 14665 14725  
 FT CHAIN 14725 14785  
 FT DOMAIN 14725 14785  
 FT CHAIN 14785 14845  
 FT DOMAIN 14785 14845  
 FT CHAIN 14845 14905  
 FT DOMAIN 14845 14905  
 FT CHAIN 14905 14965  
 FT DOMAIN 14905 14965  
 FT CHAIN 14965 15025  
 FT DOMAIN 14965 15025  
 FT CHAIN 15025 15085  
 FT DOMAIN 15025 15085  
 FT CHAIN 15085 15145  
 FT DOMAIN 15085 15145  
 FT CHAIN 15145 15205  
 FT DOMAIN 15145 15205  
 FT CHAIN 15205 15265  
 FT DOMAIN 15205 15265  
 FT CHAIN 15265 15325  
 FT DOMAIN 15265 15325  
 FT CHAIN 15325 15385  
 FT DOMAIN 15325 15385  
 FT CHAIN 15385 15445  
 FT DOMAIN 15385 15445  
 FT CHAIN 15445 15505  
 FT DOMAIN 15445 15505  
 FT CHAIN 15505 15565  
 FT DOMAIN 15505 15565  
 FT CHAIN 15565 15625  
 FT DOMAIN 15565 15625  
 FT CHAIN 15625 15685  
 FT DOMAIN 15625 15685  
 FT CHAIN 15685 15745  
 FT DOMAIN 15685 15745  
 FT CHAIN 15745 15805  
 FT DOMAIN 15745 15805  
 FT CHAIN 15805 15865  
 FT DOMAIN 15805 15865  
 FT CHAIN 15865 15925  
 FT DOMAIN 15865 15925  
 FT CHAIN 15925 15985  
 FT DOMAIN 15925 15985  
 FT CHAIN 15985 16045  
 FT DOMAIN 15985 16045  
 FT CHAIN 16045 16105  
 FT DOMAIN 16045 16105  
 FT CHAIN 16105 16165  
 FT DOMAIN 16105 16165  
 FT CHAIN 16165 16225  
 FT DOMAIN 16165 16225  
 FT CHAIN 16225 16285  
 FT DOMAIN 16225 16285  
 FT CHAIN 16285 16345  
 FT DOMAIN 16285 16345  
 FT CHAIN 16345 16405  
 FT DOMAIN 16345 16405  
 FT CHAIN 16405 16465  
 FT DOMAIN 16405 16465  
 FT CHAIN 16465 16525  
 FT DOMAIN 16465 16525  
 FT CHAIN 16525 16585  
 FT DOMAIN 16525 16585  
 FT CHAIN 16585 16645  
 FT DOMAIN 16585 16645  
 FT CHAIN 16645 16705  
 FT DOMAIN 16645 16705  
 FT CHAIN 16705 16765  
 FT DOMAIN 16705 16765  
 FT CHAIN 16765 16825  
 FT DOMAIN 16765 16825  
 FT CHAIN 16825 16885  
 FT DOMAIN 16825 16885  
 FT CHAIN 16885 16945  
 FT DOMAIN 16885 16945  
 FT CHAIN 16945 17005  
 FT DOMAIN 16945 17005  
 FT CHAIN 17005 17065  
 FT DOMAIN 17005 17065  
 FT CHAIN 17065 17125  
 FT DOMAIN 17065 17125  
 FT CHAIN 17125 17185  
 FT DOMAIN 17125 17185  
 FT CHAIN 17185 17245  
 FT DOMAIN 17185 17245  
 FT CHAIN 17245 17305  
 FT DOMAIN 17245 17305  
 FT CHAIN 17305 17365  
 FT DOMAIN 17305 17365  
 FT CHAIN 17365 17425  
 FT DOMAIN 17365 17425  
 FT CHAIN 17425 17485  
 FT DOMAIN 17425 17485  
 FT CHAIN 17485 17545  
 FT DOMAIN 17485 17545  
 FT CHAIN 17545 17605  
 FT DOMAIN 17545 17605  
 FT CHAIN 17605 17665  
 FT DOMAIN 17605 17665  
 FT CHAIN 17665 17725  
 FT DOMAIN 17665 17725  
 FT CHAIN 17725 17785  
 FT DOMAIN 17725 17785  
 FT CHAIN 17785 17845  
 FT DOMAIN 17785 17845  
 FT CHAIN 17845 17905  
 FT DOMAIN 17845 17905  
 FT CHAIN 17905 17965  
 FT DOMAIN 17905 17965  
 FT CHAIN 17965 18025  
 FT DOMAIN 17965 18025  
 FT CHAIN 18025 18085  
 FT DOMAIN 18025 18085  
 FT CHAIN 18085 18145  
 FT DOMAIN 18085 18145  
 FT CHAIN 18145 18205  
 FT DOMAIN 18145 18205  
 FT CHAIN 18205 18265  
 FT DOMAIN 18205 18265  
 FT CHAIN 18265 18325  
 FT DOMAIN 18265 18325  
 FT CHAIN 18325 18385  
 FT DOMAIN 18325 18385  
 FT CHAIN 18385 18445  
 FT DOMAIN 18385 18445  
 FT CHAIN 18445 18505  
 FT DOMAIN 18445 18505  
 FT CHAIN 18505 18565  
 FT DOMAIN 18505 18565  
 FT CHAIN 18565 18625  
 FT DOMAIN 18565 18625  
 FT CHAIN 18625 18685  
 FT DOMAIN 18625 18685  
 FT CHAIN 18685 18745  
 FT DOMAIN 18685 18745  
 FT CHAIN 18745 18805  
 FT DOMAIN 18745 18805  
 FT CHAIN 18805 18865  
 FT DOMAIN 18805 18865  
 FT CHAIN 18865 18925  
 FT DOMAIN 18865 18925  
 FT CHAIN 18925 18985  
 FT DOMAIN 18925 18985  
 FT CHAIN 18985 19045  
 FT DOMAIN 18985 19045  
 FT CHAIN 19045 19105  
 FT DOMAIN 19045 19105  
 FT CHAIN 19105 19165  
 FT DOMAIN 19105 19165  
 FT CHAIN 19165 19225  
 FT DOMAIN 19165 19225  
 FT CHAIN 19225 19285  
 FT DOMAIN 19225 19285  
 FT CHAIN 19285 19345  
 FT DOMAIN 19285 19345  
 FT CHAIN 19345 19405  
 FT DOMAIN 19345 19405  
 FT CHAIN 19405 19465  
 FT DOMAIN 19405 19465  
 FT CHAIN 19465 19525  
 FT DOMAIN 19465 19525  
 FT CHAIN 19525 19585  
 FT DOMAIN 19525 19585  
 FT CHAIN 19585 19645  
 FT DOMAIN 19585 19645  
 FT CHAIN 19645 19705  
 FT DOMAIN 19645 19705  
 FT CHAIN 19705 19765  
 FT DOMAIN 19705 19765  
 FT CHAIN 19765 19825  
 FT DOMAIN 19765 19825  
 FT CHAIN 19825 19885  
 FT DOMAIN 19825 19885  
 FT CHAIN 19885 19945  
 FT DOMAIN 19885 19945  
 FT CHAIN 19945 20005  
 FT DOMAIN 19945 20005  
 FT CHAIN 20005 20065  
 FT DOMAIN 20005



Score:	770.00	Matches:	163
Percent Similarity:	55.71%	Conservative:	71
Best Local Similarity:	38.81%	Mismatches:	150
Query Match:	32.95%	Indels:	36
DB:	1	Gaps:	10
US-09-763-153-1 (1-1245) x FA7_MOUSE (1-446)			
QY	1	GCCAACTCCTTCCTGGAGGAGCTCCGTCACAGCAGCCTGGAGCGGAGTGCATACAGGAG	60
Db	42	AlaAsnSerLeuLeuGluLeuTrpProGlySerLeuGluArgIgLcYcAsnGluGlu	61
QY	61	ATCTGTGACTTCAGAGGCCCAAGAAATTTCCAAAATGTGGATGACACACTGCGCCTTC	12
Db	62	GlnCysSerPheGluAlaArgGluIlePheLysSerProGluArgThrLysGlnPhe	81
QY	121	TGTTCCAAAGCAGCTGCAGCGTGCACAGTGTCTGGTCTTGCCCTTGGACACCCCGTGGCC	18
Db	82	TrpIleValTyrSerAspGlyAspGlnCysAlaSerAsnPro	95
QY	181	AGCCTGTGCTGGCGGCACGCGCTTCTGCGAGCGCAGGTGAGC---	240
Db	96	-----CysGlnAsnGlyThrCysGlnAspHisLeuLysSerTyrValCysPheCys	113
QY	241	CGAGCGCTGGAGGGCGCTTCTGCGAGCGCAGGTGAGC---	297
Db	114	LeuLeuAspPheGluGlyArgAsnGlySerLysAsnGluGlnLeuLecysAla	133
QY	298	CTGGACAACGGCGCTGCAGCATTACTGCTAGAGAGTGGGCTGGCGCGC---	354
Db	134	AsnGluAsnGlyCysAspGlnTyrCysArgAspHisValGlyThrLysArgThrCys	153
QY	355	AGCTGTGGCCTGGCTACAACTGGGGACACCTCTGCAGTGTCAACCCCGCAGTGAAG	414
Db	154	SerCysHisGluAspTyrThrLeuGlnProAspGluValSerCysLysProLysValGlu	173
QY	415	TTCCCTTGGGAGCCCTGGAACGGATGGAAGAGCGCATCCTGCAAGTGTCAACCCGAGAC	474
Db	174	TyrProCysGlyArg---IleProValValGluLysArgAsnSerSerArgGlnGly	192
QY	475	ACAGAAGACCAAGAAGACCAAGTAGATCCGGCTATTGATCGGAAGATGACCGGCGG	534
Db	193	-----ArgIleValGlyAsnValCysProLys	202
QY	535	GGAGACGCCCTGGCAGGTGGTCTGCTGGACTCAAAGAAGAGTGCCTGCGGGGCA	594
Db	203	GlyLcYcProTrpGlnAlaVal---LeuLysIleAsnGlyLeuLeuLcYcGlyAla	221
QY	595	GTGTCATCCACCCCTCTCGGTGTGTACAGCGGCCCATGCTGCATGCATGAGTCCAAAG	651
Db	222	ValLeuLeuAspAlaArgTrpIleValThrAlaAlaHisCysPheAspAsnIleArgTyr	241
QY	652	-----AAGTCCTTGTCAAGCTTGGAGAGTATGACCTCGCGCTGGGAAGTGGGAG	705
Db	242	TrpGlyAsnIleThrValValMetGlyGluHisAspPheSerGluLysAspGlyAspGlu	261
QY	706	CTGGACCTGGACATCAAGGAGGTCTTCGTCACCCCACTACAGCAAGACACCACGAC	765
Db	262	GlnValArgValThrGlnValIleMetProAspLysTyrIleArgGlyLysIleAsn	281
QY	766	AATGACATCCACTGTGCACCTGGCCCGCCAGCCGACCTCTCAGACCATAGTGC	825
Db	282	HisAspIleAlaLeuLeuArgLeuHisArgProValPheThrAspTyrValValPro	301
QY	826	ATCTGCTCCCGACACGGCCTTCAGAGCGCGAGCTCAATCATCGCGCGCCAGGAC	885
Db	302	LeuCysLeuProGlnLysSerPheSerGluAsnThrLeuAlaArgIle---ArgPheSer	320
QY	886	CTCGTGACGGCTGGGCTTACCACAGCAGCGGAGAGGAGGCGCAAGAGAACCGGACC	945
Db	321	ArgValSerGlyTrpGlyGlnLeuLeuAspArgGlyAlaThrAla	335
QY	946	TTCGTCTCAACTTCATCAAGATTCCCTGGTCCCGCAATAGTGCACGAGGTCA	1005



DISULFID	57	62	BY SIMILARITY.
DISULFID	90	101	BY SIMILARITY.
DISULFID	95	110	BY SIMILARITY.
DISULFID	112	121	BY SIMILARITY.
DISULFID	127	138	BY SIMILARITY.
DISULFID	134	148	BY SIMILARITY.
DISULFID	150	163	BY SIMILARITY.
DISULFID	197	207	N-LINKED (GLCNAC. . .) (POTENTIAL)
CARBOHYD	197	207	N-LINKED (GLCNAC. . .) (POTENTIAL)
CARBOHYD	297	297	N-LINKED (GLCNAC. . .) (POTENTIAL)
ACT. SITE	306	306	CHARGE RELAY SYSTEM.
ACT. SITE	402	402	CHARGE RELAY SYSTEM.
ACT. SITE	402	402	CHARGE RELAY SYSTEM.
ACT. SITE	402	402	CHARGE RELAY SYSTEM.
VARIANT	418	418	G -> E (IN HEMOPHILIA B).
VARIANT	418	418	1P6537C46A960ED CRC64.
VARIANT	418	418	418927 ver.

	InterPro:	IPR002383; GLA_blood.		Alignment Scores:	1_5le-33	Length:	452
R	InterPro:	IPR001254; Ser_protease_Try.		Pred. No.: 763.00	Matches:	Conservative:	158
R	pfam:	PF00008; EGF_2.		Percent Similarity: 53.97%	Mismatches:	: 73	
R	pfam:	PF00089; trypsin; 1.		Best Local Similarity: 36.92%	Gaps:	:	141
R	pfam:	PF00594; gla; 1.		Query Match: 32.65%	Indels:	:	56
R	PRINTS:	PR00722; CHYMOTRYPSIN.		DB: 1			12
R	PRINTS:	PR00001; GLABLOOD.					
R	SMART:	SM00179; EGF_CA; 1.		US-09-763-153-1 (1-1245) x FA9_CANFA (1-452)			
R	SMART:	SM00069; GLA; 1.					
R	PROSITE:	PSM0020; Tryp_Spc; 1.					
R	PROSITE:	PSM0010; ASX_HYDROXYL; 1.					
R	PROSITE:	PSM0022; EGF_1; 1.					
R	PROSITE:	PSM1186; EGF_2; 2.					
R	PROSITE:	PSM1187; EGF_CA; 1.					
R	PROSITE:	PSM0011; GLUT_CARBOXYLATION; 1.					
R	PROSITE:	PSM0240; TRYPSIN_DOM; 1.					
R	PROSITE:	PSM0134; TRYPSIN_SER; 1.					
R	PROSITE:	PSM0135; TRYPSIN_HIS; 1.					
W	Blood coagulation;	Plasma; Serine protease;					
W	Hyaluronate synthetase-like domain	Vitamin K; Gamma-carboxyglutamic acid;					
W	Membrane-type I tyrosine kinase	Zymogen; Signal; EGF-like domain;					
W	Repeat; Disease mutation.						
T	SIGNAL	BY SIMILARITY.					
T	PROPPEP	21					
T	CHAIN	22					
T	CHAIN	40					
T	PROPPEP	184					
T	CHAIN	218					
T	DOMAIN	86					
T	DOMAIN	123					
T	DOMAIN	164					
T	DOMAIN	218					
T	SITE	183					
T	SITE	217					
T	MOD_RES	46					
F	MOD_RES	47					
F	MOD_RES	54					
F	MOD_RES	56					
F	MOD_RES	59					
F	MOD_RES	60					
F	MOD_RES	65					
F	MOD_RES	66					
F	MOD_RES	69					
F	MOD_RES	72					
F	MOD_RES	75					
F	MOD_RES	79					
F	MOD_RES	103					
F	MOD_RES	103					





Db 221 PheAsnAspPheThrArgValValGlyGlyAlaValAlaLysProGlyGlnPheProTrp 240  
 QY 550 CAGGTGCTGCTGGACTCAAGAGAGAGCTGGCTGGGGGAGTGTCTATCCACCCC 609  
 Db 241 GlnValVal---LeuAsnGlyLysValAlaPheCysGlySerIleValAsnGlu 259  
 QY 610 TCCTGGGTGTGACAGCGGCCACTGCATGGATGAGTCAACAAGCTCTGTCTGAGGCTT 669  
 Db 260 LysTrpIleValThrAlaAlaHisCysValGluThrGlyValLysIleThrValAla 279  
 QY 670 GGAGATGATGACCTCGCGCGCTGGGAGAGTGGAGCTGGACCTGGACATCAAGGAGGTC 729  
 Db 280 GlyGluHisAsnIleGluThrGluHisThrGluGlnLysArgAsnValIleArgIle 299  
 QY 730 TTCGTCCACCCCACTACAGCAGCAGCACC-----GACATGACATCGCTGCTGCT 783  
 Db 300 IleProHisAsnTyrAsnAlaIleAsnLysTyrAsnHisAspIleAlaLeu 319  
 QY 784 CACCTGGCCCGCCCGCCCTCTCGCAGCACCATGAGTCCCTGCTCCCGCAGCAGC 843  
 Db 320 GluLeuAspGluProLeuValLeuAsnSerTyrValThrProIleCysIleAlaAspLys 339  
 QY 844 GGCCTTGACAGCGGAGCTCAATCAGCGCGGAGAGACCTCGTGCAGGGCTGGGC 903  
 Db 340 GluTyrThrAsnIlePheLeuLysPheGly-----SerGlyTyrValSerGlyTrpGly 357  
 QY 904 -----TACCACAGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957  
 Db 358 ArgValPheHisLysGlyArgSer-----AlaLeuValLeuGln 370  
 QY 958 TTCATCAAGATTCCTGCTGCTCCCGCACATCAGTGCAGCGAGTGCATGAGCAACATGGTG 1017  
 Db 371 TyrLeuArgValProLeuValAlaPheArgAlaThrCysLeuArgSerThrLysPheThrIle 390  
 QY 1018 TCTGAGACATGCTGTGTGGGGCATCTCGGGGACCGGAGTGCCTGCTGGTGGCTGTGAGC 1077  
 Db 391 TyrAsnAsnMetPheCysAlaGlyPheHisGluGlyArgAspSerCysGlnGlyAsp 410  
 QY 1078 AGTGGGGGGCCATGCTGCTCTCTCCAGCGCAGCTGCTGTGCTGGCTGTGAGC 1137  
 Db 411 SerGlyGlyProHisValThrGluValGluGlyThrSerPheLeuThrGlyIleIleSer 430  
 QY 1138 TGGGGTGGGCTGTGGCTCTTCAACTAGCGGCTTTACACAAAGTACAGCGCTAC 1197  
 Db 431 TrpGlyGluGlyCysAlaMetLysGlyLysTyrGlyIleThrLysValSerArgTyr 450  
 QY 1198 CTCGACTGGATC 1209  
 Db 451 ValAsnTrpIle 454  
 RESULT 18  
 FA9\_MOUSE STANDARD; PRT; 459 AA.  
 P16294;  
 01-AUG-1990 (Rel. 15, Created)  
 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Coagulation factor IX precursor (EC 3.4.21.22) (Christmas factor)  
 DE (Fragment).  
 GN F9 OR CF9.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90215309; PubMed=2323576;  
 RA Wu S.-M., Stafford D.W., Ware J.;  
 RT "Deduced amino acid sequence of mouse blood-coagulation factor IX.";  
 RL Gene 86:275-278(1990).  
 RN [2]  
 RP SEQUENCE OF 168-451 FROM N.A.  
 RX MEDLINE=90152675; PubMed=2303254;

RA Sarker G., Koerberl D.D., Sommer S.S.;  
 RT "Direct sequencing of the activation peptide and the catalytic domain  
 of the factor IX gene in six species.";  
 RL Genomics 6:133-143(1990).  
 CC -1- FUNCTION: FACTOR IX IS A VITAMIN K-DEPENDENT PLASMA PROTEIN THAT  
 PARTICIPATES IN THE INTRINSIC PATHWAY OF BLOOD COAGULATION BY  
 CONVERTING FACTOR X TO ITS ACTIVE FORM IN THE PRESENCE OF CA++  
 IONS, PHOSPHOLIPIDS, AND FACTOR VIIIa.  
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes one Arg-|-Ile bond in factor X to  
 form factor Xa.  
 CC -1- SUBUNIT: FACTOR IX IS ACTIVATED BY FACTOR XIA, WHICH EXCISES THE  
 ACTIVATION PEPTIDE AND THUS PRODUCES A MOLECULE CONSISTING OF 2  
 CHAINS HELD TOGETHER BY ONE OR MORE DISULFIDE BONDS.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; M23109; AAA37629.1; -;  
 DR EMBL; M26236; AAA37630.1; -;  
 DR PIR; JQ0419; JQ0419.  
 DR HSSP; P00740; 1CFH.  
 DR MEROPS; S01.214; -;  
 DR MGD; MGI-88384; F9.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR000742; EGF-2.  
 DR InterPro; IPR001881; EGF-Ca.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR InterPro; IPR000294; VitK\_dep\_GLA.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF00089; trypsin; 1.  
 DR Pfam; PF00594; g1a; 1.  
 DR SMART; SM00179; EGF\_CA; 1.  
 DR SMART; SM00069; GLA; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 2.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 DR PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Blood coagulation; Plasma; Serine protease; Calcium-binding;  
 KW Hydrolase; Glycoprotein; Vitamin K; Gamma-carboxyglutamic acid;  
 KW Hemophilia; Hydroxylation; Zymogen; Signal; EGF-like domain;  
 KW Repeat.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 16  
 FT PROPEP 17 34  
 FT CHAIN 35 180  
 FT PROPEP 181 224  
 FT CHAIN 225 459  
 FT DOMAIN 81 117  
 FT DOMAIN 118 159  
 FT DOMAIN 225 459  
 FT SITE 180 181  
 FT SITE 224 225  
 FT MOD\_RES 41 41  
 FT MOD\_RES 42 42  
 FT MOD\_RES 49 49  
 FT MOD\_RES 51 51  
 FT  
 BY SIMILARITY.  
 FACTOR IXA (ACTIVE FORM) LIGHT CHAIN.  
 ACTIVATION PEPTIDE.  
 FACTOR IXA (ACTIVE FORM) HEAVY CHAIN.  
 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).  
 EGF-LIKE 2.  
 SERINE PROTEASE.  
 CLEAVAGE (BY FACTOR XIA) (BY SIMILARITY).  
 CLEAVAGE (BY FACTOR XIA) (BY SIMILARITY).  
 GAMMA-CARBOXYGLUTAMIC ACID  
 (BY SIMILARITY).  
 GAMMA-CARBOXYGLUTAMIC ACID  
 (BY SIMILARITY).  
 GAMMA-CARBOXYGLUTAMIC ACID  
 (BY SIMILARITY).  
 GAMMA-CARBOXYGLUTAMIC ACID  
 (BY SIMILARITY).

FT MOD\_RES 54 54 GAMMA-CARBOXYGLUTAMIC ACID  
 FT (BY SIMILARITY).  
 FT MOD\_RES 55 55 GAMMA-CARBOXYGLUTAMIC ACID  
 FT (BY SIMILARITY).  
 FT MOD\_RES 60 60 GAMMA-CARBOXYGLUTAMIC ACID  
 FT (BY SIMILARITY).  
 FT MOD\_RES 61 61 GAMMA-CARBOXYGLUTAMIC ACID  
 FT (BY SIMILARITY).  
 FT MOD\_RES 64 64 GAMMA-CARBOXYGLUTAMIC ACID  
 FT (BY SIMILARITY).  
 FT MOD\_RES 67 67 GAMMA-CARBOXYGLUTAMIC ACID  
 FT (BY SIMILARITY).  
 FT MOD\_RES 70 70 GAMMA-CARBOXYGLUTAMIC ACID  
 FT (BY SIMILARITY).  
 FT MOD\_RES 74 74 GAMMA-CARBOXYGLUTAMIC ACID  
 FT (BY SIMILARITY).  
 MOD\_RES 98 98 HYDROXYLATION (BY SIMILARITY).  
 ACT\_SITE 265 265 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 313 313 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 409 409 CHARGE RELAY SYSTEM.  
 FT DISULFID 52 57 BY SIMILARITY.  
 FT DISULFID 85 96 BY SIMILARITY.  
 FT DISULFID 90 105 BY SIMILARITY.  
 FT DISULFID 107 116 BY SIMILARITY.  
 FT DISULFID 122 133 BY SIMILARITY.  
 FT DISULFID 129 143 BY SIMILARITY.  
 FT DISULFID 145 158 BY SIMILARITY.  
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 211 211 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 363 363 H -> Q (IN REF. 2).  
 FT CONFLICT 388 388 T -> I (IN REF. 2).  
 SQ SEQUENCE 459 AA; 51635 MW; EF439C840D6BC1A CRC64;

Alignment Scores:  
 Pred. No.: 1,36e-31 Length: 459  
 Score: 726.00 Matches: 155  
 Percent Similarity: 50.91% Conservative: 68  
 Best Local Similarity: 35.39% Mismatches: 151  
 Query Match: 31.07% Indels: 64  
 DB: 1 Gaps: 12

US-09-763-153-1 (1-1245) x FA9\_MOUSE (1-459)

QY 13 CTGAGAGGCTCCGTCACAGCAGCCCTGGAGCGGAGTGCATGAGAGAGATCTGTGACTTC 72  
 |||||  
 40 LeuGluGluPheValArgGlyAsnLeuGluArgGluCysIleGluArgCysSerPhe 59  
 |||||  
 73 GAGGAGCCAGGAATTTCCAAATGTGGATGACACACTGGCCTTCTGGTCCCAAGCAC 132  
 |||||  
 60 GluGluAlaArgGluValPheGluAsnThrGluLysThrGluPheTrpLysGlnTyr 79  
 |||||  
 133 GTCAGCGGTGACAGTCTGGTCTTGGCCCTTGGAGCACCCTGCGCCAGCCTGTGCTGC 192  
 |||||  
 80 ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu 91  
 |||||  
 193 GGGCAGCGCAGTGCATCGACGCGATCGGAGTGCATGTCGACGTGCGCGAGCGCTGG 252  
 |||||  
 92 AsnGlyGlyIleCysLysAspAspIleSerSerTyrGluCysTrpCysGlnValGlyPhe 111  
 |||||  
 253 GAGGCGCGCTTCTCCAGCGGAGGTGAGTCTCTCAATGTCGCTGGGACACGCGCGC 312  
 |||||  
 112 GluGlyArgAsnCysGluLeuAspAlaThr-----CysAsnIleLysAsnGlyArg 128  
 |||||  
 313 TGCAGCATTTACTCGTAGAGAGGTGGCTGGCGG---CGCTGTAGCTGTGGCGCTGGC 369  
 |||||  
 129 CysLysGlnPheCysLysAsnSerProAspAsnLysValIleCysSerCysThrGluGly 148  
 |||||  
 370 TACAAGCTGGGAGCAGCTCTCGAGTGTACCCCGCAGTGAAGTCCCTTGTGGAGG 429  
 |||||  
 149 TyrGlnLeuAlaGluAspGlnLysSerCysGluProThrValProPheProCysGlyArg 168  
 |||||  
 430 CCC-----TGGAGCGGATG 444

Db 169 AlaSerIleSerTyrSerSerLysLysIleThrArgAlaGluThrValPheSerAsnMet 188  
 |||||  
 QY 445 GAGAAGAGCGCAGTCACTAC-----  
 |||||  
 Db 189 AspTyrGluAsnSerThrGluAlaValPheIleGlnAspAspIleThrAspGlyAlaIle 208  
 |||||  
 QY 463 CTGAACGAGACACAGACAGACCAAGAACCAAGTAGAT--CGCGGCTCAATTGATGGG 519  
 |||||  
 Db 209 LeuAsnAsnValThrGluSerSerGluSerLeuAsnAspPheThrArgValValGlyGly 228  
 |||||  
 QY 520 AAGATGACACGCGGGGAGACAGCCCTGCGAGGTGCTCTGCTGGACTCAAGAGAGAAG 579  
 |||||  
 Db 229 GluAsnAlaLysProGlyGlnIleProTrpGlnValIle--LeuAsnGlyGluIleGlu 247  
 |||||  
 QY 580 CTGGCTGCGGGGAGTCTCTATCCACCCCTCTGGTGGTGTGACAGCGGCCCTGCTGATG 639  
 |||||  
 Db 248 AlaPheCysGlyGlyAlaIleAsnGluLysTrpIleValThrAlaAlaHisCysLeu 267  
 |||||  
 QY 640 GATGACTCAAGAACCTCTTGTGAGGTGAGAGATGACCTGCGCGCTGGGAGAAG 699  
 |||||  
 Db 268 LysProGlyAspLysIleGluValAlaGluLysTrpIleValThrAsnIleAspLysLysGluAsp 287  
 |||||  
 QY 700 TGGGAGCTGGACCTGGACATCAAGGAGTCTCTGCTCCACCCCACTACAGCAAGAGACCC 759  
 |||||  
 Db 288 ThrGluGlnArgArgAsnValIleArgThrIleProHisGlnTyrAsnAlaThrIle 307  
 |||||  
 QY 760 ACC-----GACATGACATCCACTGCTGACCTGCGCCAGCCGCGCTGGGAGAG 813  
 |||||  
 Db 308 AsnLysTyrSerHisAspIleAlaLeuGluLeuLysProLeuIleLysSer 327  
 |||||  
 QY 814 ACCATAGTCCCATCTGCTCCCGGACAGCGGCTTGCAGAGCGGAGCTCAATCAG--- 870  
 |||||  
 Db 328 TyrValThrProLysCys-----ValAlaAsnArgGluTyrThrAsnIle 342  
 |||||  
 QY 871 -----GCCGCGCAGGAGACCCCTGCTGAGCGGCTGGGCTTACCACAGCAGC 915  
 |||||  
 Db 343 PheLeuLysPheGlySerGly-----TyrValSerGlyTrpGlyLysValPheAsn 359  
 |||||  
 QY 916 CGAGAGAGGAGGCGCAAGAGAACCGACCTCTGCTCAACTTCTCAAGATTCCTCCGTG 975  
 |||||  
 Db 360 LysGlyArgHisAlaSer-----IleLeuGlnTyrLeuArgValProLeu 374  
 |||||  
 QY 976 GTCCCGCACAAATGAGTGCAGCGAGGTCTATGACCAACATGGTGTCTGAGAACATGTGTGT 1035  
 |||||  
 Db 375 ValAspArgAlaThrCysLeuArgSerThrThrPheThrThrTyrAsnAsnMetPheCys 394  
 |||||  
 QY 1036 GCGGCGCATCTCGGGACCGCGCAGGATGCTGCGAGGCGGACAGTGGGGGGCCCATGTC 1095  
 |||||  
 Db 395 AlaGlyTyrArgGluGlyLysAspSerCysGluGlyAspSerGlyGlyProHisVal 414  
 |||||  
 QY 1096 GCCTCTTCCACGCGCAGCTGGTCTGTTGGTGGCTGCTGAGTGGGGTGGGCTGTGGG 1155  
 |||||  
 Db 415 ThrGluValGluGlyThrSerPheLeuThrGlyIleIleSerTrpGlyGluGluCysAla 434  
 |||||  
 QY 1156 CTCCTTCCAACTACGCGGTTTACACCAAGTACGCGCTTACCTGACCTGGATC 1209  
 |||||  
 Db 435 MetLysGlyLysTyrGlyIleTyrThrLysValSerArgTyrValAsnTrpIle 452  
 |||||  
 RESULT 19  
 FA10\_TROCA STANDARD; PRT; 376 AA.  
 ID FA10\_TROCA  
 AC P81428;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Coagulation factor X (EC 3.4.21.6) (Trocarin prothrombin activator).  
 OS Tropidochelis carinata (Australian rough-scaled snake).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidae;  
 OC Elapidae; Notechidae; Tropidochelis.  
 OX NCBI\_TaxID=100989;  
 RN [1]



QY 721 AAGGAGGTCTTCCTCCAC-----CCCAACTAC----- 747  
 Db 201 AsplylletyrrValHisThrLysPheValProProAsnTyrTyrValHisGlnAsn 220  
 QY 748 ---AGCAAGAGCACCACGACATGACATGCGACTGCTGCACCTGCGCCAGCCGCCACCC 804  
 Db 221 PheAspArgValAlaTyrAspTyrAspPheAlaAlaIleArgMetLysThrProIleGln 240  
 QY 805 CTCTCCGACACCATAGTCCCATTCCTCCGCGACAGGGCTTCAGAGCGCGAGCTC 864  
 Db 241 PheSerGluAsnValValProAlaCysLeuProThrAlaAspPheAla---AsnGluVal 259  
 QY 865 AATCAGCGCGCGCAGGAGACCTCGTACGGCGCTGCGGTACACAGCAGCGCAGAGAG 924  
 Db 260 LeuMetLysGlnAspSerGlyIleValSerGlyPheGly-----ArgIle 274  
 QY 925 GAGCGCAAGAGAAACCGACCTTCCTCTCACTCATCAAGATTCCTGCTGCTCCGCGCAG 984  
 Db 275 GlnPheLysGlnProThrSerAsnThrLeuLysValIleThrValProTyrValAspArg 294  
 QY 985 AATGAGTGCAGGAGTGCATGAGCAACATGCTGTGAGAACATGCTGTGCGCGGCATC 1044  
 Db 295 HisThrCysMetLeuSerSerAspPheArgIleThrGlnAsnMetPheCysAlaGlyTyr 314  
 QY 1045 CTCGGGCGCGCAGGATCCCTGCGAGGCGACAGTGGGGCGCCATGCTGCTCCTTC 1104  
 Db 315 AspThrLeuProGlnAspAlaCysGlnGlyAspSerGlyProHisIleThrAlaTyr 334  
 QY 1105 CACGCGACCTGTTCTCTGCTGGCGCTGTGAGCTGGGTGAGGCTGTGGGCTCTTCAC 1164  
 Db 335 ArgAspThrHisPheIleThrGlyIleIleSerTyrGlyGluGlyCysAlaArgLysGly 354  
 QY 1165 ACTACGGCGTTTACACCAAGTACCGCTACCTACCTGCTGAGTGCATC 1209  
 Db 355 LysTyrGlyValTyrThrLysValSerLysPheIleProTyrPile 369

RESULT 20  
 PRTC\_CANFA  
 ID PRTC\_CANFA STANDARD; PRT; 157 AA.  
 AC Q28278;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Vitamin-K dependent protein C (EC 3.4.21.69) (Autoprothrombin IIA)  
 SE (Anticoagulant protein C) (Blood coagulation factor XIV) (Fragment).  
 PROC.  
 Canis familiaris (Dog).  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 NCBI\_TaxID=9615;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94318474; PubMed=8043441;  
 RA Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;  
 RT "A comparative study of partial primary structures of the catalytic  
 region of mammalian protein C.";  
 RL Br. J. Haematol. 86:590-600(1994).  
 CC -!- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT  
 CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA  
 CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.  
 CC -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va  
 CC and VIIIA.  
 CC -!- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC EMBL; D43751; BAA07808.1; -.  
 DR HSP; P04070; LPCU.  
 DR MEROPS; S01.218; -.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; PARTIAL.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Blood coagulation; Glycoprotein; Serine protease; Hydrolase.  
 FT NON\_TER 1 1  
 FT ACT\_SITE 26 26 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 125 125 CHARGE RELAY SYSTEM.  
 FT DISULFID 96 110 BY SIMILARITY.  
 FT DISULFID 121 149 BY SIMILARITY.  
 FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT NON\_TER 157 157  
 SQ SEQUENCE 157 AA; 17262 MW; E8B1BACF49220DFB CRC64;

Alignment Scores:  
 Pred. No.: 4.19e-31 Length: 157  
 Score: 717.00 Matches: 132  
 Percent Similarity: 86.96% Conservative: 8  
 Best Local Similarity: 81.96% Mismatches: 17  
 Query Match: 30.68% Indels: 4  
 DB: 1 Gaps: 1

US-09-763-153-1 (1-1245) x PRTC\_CANFA (1-157)

QY 694 GAGAAGTGGGAGCTGGACCTGGACATCAAGAGAGCTTGTCTCCACCCCACTACAGCAAG 753  
 Db 1 GlulysGlyGluMetAspValAspIleLysGluValLeuIleHisProAsnTyrSerLys 20  
 QY 754 AGCACCCAGCAATGACATCGCACTGCTGCACCTGCCAGCCGCCAGCCGCCCTCTCCGAG 813  
 Db 21 SerThrThrAspAsnAspIleAlaLeuLeuHisLeuAlaGlnProAlaIlePheSerGln 40  
 QY 814 ACCATAGTGGCCATCTCCCTCCCGACAGCGGCTTGCAGAGCGGAGCTCAATCAGGCC 873  
 Db 41 ThrIleValProIleCysLeuProAspSerGlyLeuAlaGluArgGluLeuThrGlnVal 60  
 QY 874 GGCAGAGAGCCCTCTGACGGGCTGGGCTACACAGACCGCAGAGAGAGGAGCCCAAG 933  
 Db 61 GlyGlnGluThrValValThrGlyTyrGlyTyrArgSer-----GluThrLys 76  
 QY 934 AGAACCCGACCTCTGCTCAACTTCATCAAGATTCCCGTGTCCGCGCACATGAGTGC 993  
 Db 77 ArgAsnArgThrPheValLeuAsnPheIleAsnIleProValAlaProHisAsnGluCys 96  
 QY 994 AGCAGGTGATGAGCAACATGGTGTGTGAGAACATGCTGTGCGGGCATCTCTCGGGAC 1053  
 Db 97 IleGlnAlaMetTyrAsnMetIleSerGluAsnMetLeuCysAlaGlyIleLeuGlyAsp 116  
 QY 1054 CGCAGAGATCCCTGCGAGGCGGACAGTGGGGGCCCATGCTGCGCTCTTCCAGGCGACC 1113  
 Db 117 SerArgAspAlaCysGluGlyAspSerGlyTyrGlyProMetValThrSerPheArgGlyThr 136  
 QY 1114 TGGTCTCTGCTGGGCGCTGGTGGCTGGGCTGTGGCTGTGCTTCACTACACTAGCGC 1173  
 Db 137 TrpPheLeuValGlyLeuValSerTrpGlyGlyGlyCysGlyArgLeuHisAsnTyrGly 156  
 QY 1174 GTT 1176  
 Db 157 Ile 157

Search completed: March 17, 2003, 17:32:35  
 Job time : 57 secs

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: March 17, 2003, 17:24:24 ; Search time 73.5 Seconds  
(without alignments)  
6980.372 Million cell updates/sec

Title: US-09-763-153-1

Perfect score: 2337

Sequence: 1 gccactctcttcggagga.....acaaggaagccccccagaag 1245

Bring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 60 summaries

Command line parameters:

-MODEL=frame-n2p.model -DEV=xlp  
-Q/cgn2\_l1/USPTO\_spool/US09763153/runat\_11032003\_084247\_1993/app\_query.fasta.1.1415  
-DB-SPTRMBL\_21 -QMT-fastan -SUFFIX=n2p.rspt -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi  
-LIST=60 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=20  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09763153 @cgn.1.1.86 @runat\_11032003\_084247\_1993 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THRAIDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_Organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	1890.5	80.9	456	6	Q9TTR0		Q9Ttr0 canis famil

2	1631	69.8	460	11	Q91WN8	Q91wn8 mus musculu
3	1625	69.5	460	11	Q99PC6	Q99pc6 mus musculu
4	818.5	35.0	482	11	Q63207	Q63207 rattus norv
5	794	34.0	481	11	Q88947	Q88947 mus musculu
6	793	33.9	481	11	O54740	O54740 mus musculu
7	793	33.9	481	11	Q99L32	Q99l32 mus musculu
8	787	33.7	433	13	Q90YK1	Q90yK1 brachydanio
9	781	33.4	701	4	Q96PQ8	Q96pQ8 homo sapien
10	773	33.1	469	6	Q9GMD9	Q9gmd9 ornithorhyn
11	772	33.0	446	11	Q61109	Q61109 mus musculu
12	740	31.7	461	6	Q95ND7	Q95nd7 pan troglod
13	736	31.5	456	4	Q14316	Q14316 homo sapien
14	718	30.7	461	6	Q95ND6	Q95nd6 pan troglod
15	525	22.5	608	13	Q9PTW7	Q9ptW7 struthio ca
16	512	21.9	607	13	Q91001	Q91001 gallus gall
17	510.5	21.8	799	11	Q9DB10	Q9db10 mus musculu
18	477.5	20.4	653	11	Q8VCS4	Q8vcs4 mus musculu
19	470	20.1	399	11	Q9CQW3	Q9cqW3 mus musculu
20	440.5	18.8	1379	5	Q9V4N6	Q9v4N6 drosophila
21	433	18.5	624	11	Q9DAT3	Q9dat3 mus musculu
22	431.5	18.5	685	11	Q91WP0	Q91wp0 mus musculu
23	431	18.4	387	5	Q9XV57	Q9xy57 ctenocephal
24	427.5	18.3	420	13	Q90504	Q90504 eptatretus
25	427	18.3	624	11	Q91V47	Q91v47 mus musculu
26	426	18.2	685	11	Q92338	Q92338 mus musculu
27	424.5	18.2	701	11	Q9JJS9	Q9jjs9 rattus norv
28	422	18.1	767	13	Q9DGR2	Q9dgr2 xenopus lae
29	420	18.0	638	11	Q8ROP5	Q8rop5 mus musculu
30	420	18.0	678	11	Q9JJS8	Q9jjs8 rattus norv
31	419.5	18.0	277	5	Q96899	Q96899 scolopendra
32	418.5	17.9	855	11	Q9JJI7	Q9jji7 rattus norv
33	417.5	17.9	643	6	Q97506	Q97506 sus scrofa
34	417.5	17.9	812	11	Q91WJ5	Q91wj5 mus musculu
35	417	17.8	573	5	Q9V516	Q9v516 drosophila
36	416	17.8	1186	5	Q9VSU2	Q9vsu2 drosophila
37	416	17.8	1449	5	Q9UII2	Q9uli2 drosophila
38	416	17.8	1462	5	Q9UII3	Q9uli3 drosophila
39	416	17.8	2382	5	Q9B119	Q9bli9 drosophila
40	416	17.8	2409	5	Q960G6	Q960g6 drosophila
41	415.5	17.8	453	11	Q8VDE0	Q8vde0 mus musculu
42	415	17.8	541	11	Q9QX90	Q9qx90 rattus norv
43	415	17.8	623	11	Q9JJP3	Q9jjp3 rattus norv
44	415	17.8	643	11	Q9QX84	Q9qx84 rattus norv
45	414	17.7	435	5	Q9NEY2	Q9nfY2 anopheles g
46	413.5	17.7	366	11	Q9QX85	Q9qx85 rattus norv
47	413	17.7	490	11	Q920K3	Q920k3 rattus norv
48	411	17.6	810	4	Q15146	Q15146 homo sapien
49	410	17.5	560	4	Q14520	Q14520 homo sapien
50	410	17.5	787	5	Q9VEY6	Q9vey6 drosophila
51	409	17.5	467	5	Q967X8	Q967x8 panulirus a
52	409	17.5	698	13	Q9P071	Q9pu71 xenopus lae
53	408	17.5	722	13	Q9PS25	Q9ps25 lampetra ja
54	407	17.4	334	6	O46507	O46507 papio hamad
55	406.5	17.4	275	4	Q96RZ6	Q96rz6 homo sapien
56	406.5	17.4	1047	5	Q9VZH2	Q9vzh2 drosophila
57	404.5	17.3	581	5	Q960I5	Q960i5 drosophila
58	404.5	17.3	733	5	Q9VYX9	Q9vtx9 drosophila
59	404.5	17.3	1185	5	Q9VW38	Q9vw38 drosophila
60	404	17.3	246	11	Q9Z1R9	Q9z1r9 mus musculu

## ALIGNMENTS

RESULT 1  
Q9TTR0 ID Q9TTR0 PRELIMINARY; PRT; 456 AA.  
AC Q9TTR0;  
DT 01-MAY-2000 (TEMBLrel. 13, Created)  
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)  
DE 01-MAR-2002 (TEMBLrel. 20, Last annotation update)  
GN Protein C precursor.  
OS Canis familiaris (Dog).



```
RESULT 2
Q91WN8
ID Q91WN8 PRELIMINARY; PRT; 460 AA.
AC Q91WN8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to protein C.
GN PROC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
TISSUE=LIVER;
Strausberg R.;
Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013896; AAH13896.1; -.
DR MGD; MGI:97771; Proc.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001234; Ser_protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00594; gla; 1.
DR Pfam; PF00089; trypsin; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_2.
DR PROSITE; PS01187; EGF_CA; UNKNOWN_1.
DR PROSITE; PS00011; GLU CARBOXYLATION; UNKNOWN_1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 460 AA; 51818 MW; 0117F26E68FCC274 CRC64;

Alignment Scores:
Pred. No.: 1.55e-110 Length: 460
Score: 1631.00 Matches: 290
Percent Similarity: 83.74% Conservative: 55
Best Local Similarity: 70.39% Mismatches: 63
Query Match: 69.79% Indels: 4
DB: 11 Gaps: 3

us-09-763-153-1 (1-1245) x Q91WN8 (1-460)
QY 1 GCCAACTCTCTGAGGAGCTCCGTCACAGCAGCCTGGAGCGGAGTGATGATAGAGGAG 60
Db 42 AlaAsnSerPheLeuGluGluMetArgProGlySerLeuGluArgGluCysMetGluGlu 61
QY 61 ATCTGTGACTTCGAGGAGCCCAAGAAATTTCCAAATGTGGATGACACACTGGCCTTC 120
Db 62 IleCysAspPheGluGluAlaGlnGluPheGlnAsnValGluAspThrLeuAlaPhe 81
QY 121 TGGTCCAAAGCAGCTCGACGCTGACAGTGCTGTGCTTGGCCCTGGAGCAGCCCTGCGCC 180
Db 82 TrpIleLysTyrPheAspGlyAspGlnCysSerAlaProPheLeuAspHisGlnCysAsp 101
QY 181 AGCTGTGCTCGGGCAGCGCAGCTGTCATCGACGCATCGGACGCTTCAGTTCGCACTGC 240
Db 102 SerProCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysSerCys 121
QY 241 CGCAGCGCTGGAGGCGGCTTCGCGAGCGGAGTGAGCTTCCTCAATGTCGCTGCTG 300
Db 122 AspLysGlyTrpGluGlyLysPheCysGlnGlnGluLeuArgPheGlnAspCysArgVal 141
QY 301 GACACGCGGCTGTCACGCTTACTGCTAGGAGGCTGGCTGGCGGCTGTAGCTGT 360
Db 142 AsnAsnGlyGlyCysLeuHisTyrCysLeuGluGluSerAsnGlyArgArgCysAlaCys 161

us-09-763-153-1 (1-1245) x Q91WN8 (1-460)
QY 361 CGGCTGCTACAAAGCTGGGGGAGCAGCCTCTCTGAGTGTACCCCGGAGTGAAGTCCCT 420
Db 162 AlaProGlyTyrGluLeuAlaAspHisMetArgCysLysSerThrValAsnProPhe 181
QY 421 TGTGGAGGCGCTGGAAGCGGATGAGAGAGCGCAGTACCTCACTGAACAGGACACAGAA 480
Db 182 CysGlyLysLeuGlyArgTrpIleGluLysArgLysIleLeuLysArgAspThr--- 200
QY 481 GACCAAGAAGACCAAGTA-----GATCCGCGCTCATTTGATGGGAAGATGACCAAGCGG 534
Db 201 AspLeuGluAspGluLeuGluProAspProArgIleValAsnGlyThrLeuThrLysGln 220
QY 535 GGAGACAGCCCTGGCAGGTGGTCTCTGCTGAGCTCAAGAGAAAGAGCTGGCTGGGGGCA 594
Db 221 GlyAspSerProTrpGlnAlaIleLeuLeuAspSerLysLysLysLeuAlaCysGlyGly 240
QY 595 GTGCTCATCCACCCTCTCTGGTGTGACAGCGGCCCTGTCATGGATGATGATGATGATGAT 654
Db 241 ValLeuIleHisThrSerTrpValLeuThrAlaAlaHisCysValGluGlyThrLysLys 260
QY 655 CTCCTGTGAGCTGGAGAGTATGACCTGGCGCTGGAGAGTGGAGAGTGGAGCTGACCTG 714
Db 261 LeuThrValArgLeuGlyGluTyrAspLeuArgArgAspHisTrpGluLeuAspLeu 280
QY 715 GACATCAAGGAGGCTTCTGCCACCCCACTACAGCAGAGCAGCAGCAGCAGCAGCAGCATC 774
Db 281 AspIleLysGluIleLeuValHisProAspTyrThrArgSerSerAspAsnAspIle 300
QY 775 GCATGCTGACCTGGCGCCAGCCCTCTCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 834
Db 301 AlaLeuLeuArgLeuAlaGlnProAlaThrLeuSerLysThrIleValProIleCysLeu 320
QY 835 CCGGACAGCGCTGGAGAGCGGAGCTCAATAGCGCGCGGAGAGAGAGAGAGAGAGAGAGAG 894
Db 321 ProAsnAsnGlyLeuAlaGln--GluLeuThrGlnAlaGlyGlnGluThrValValThr 339
QY 895 GGCTGGGCTACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 954
Db 340 GlyTrpGlyTyrGlnSerAspArgIleLysAspGlyArgArgAsnArgThrPheIleLeu 359
QY 955 AACTTCATCAAGATTCCTGCTCCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1014
Db 360 ThrPheIleArgIleProLeuValAlaArgAsnGluCysValGluValMetLysAsnVal 379
QY 1015 GTGCTGAGACATGCTGTGCTGGGCGATCCTCGGAGCAGCGGAGAGAGAGAGAGAGAG 1074
Db 380 ValSerGluAsnMetLeuCysAlaGlyIleIleGlyAspThrArgAspAlaCysAspGly 399
QY 1075 GACAGTGGGGGCGCCATGCTGCTCTCCAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1134
Db 400 AspSerGlyGlyProMetValValPhePheArgGlyThrTrpPheLeuValGlyLeuVal 419
QY 1135 AGCTGGGCTGAGGCTGTGGGCTCTTCACTACAGCGGCTTTTACACCAAGTCAAGCCGC 1194
Db 420 SerTrpGlyGluGlyCysGlyHisThrAsnAsnTyrGlyIleTyrThrLysValGlySer 439
QY 1195 TACTTCAGCTGGATCCATGGCAGCATCAGAGACAAG 1230
Db 440 TyrLeuLysTrpIleHisSerTyrIleGlyGluLys 451

RESULT 3
Q99PC6
ID Q99PC6 PRELIMINARY; PRT; 460 AA.
AC Q99PC6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Anticoagulant protein C.
GN PROC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

RN SEQUENCE FROM N.A.  
 RP STRAIN=C57BL;  
 RC Korf 1.;  
 RA "Complete sequence of UC72A01.";  
 RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 RL -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 CC  
 DR EMBL; AF318182; AAK07918.1; -.  
 DR HSP; P04070; IPCU.  
 DR MEROPS; S01.218; -.  
 DR MGD; MGI:97771; Proc.  
 DR InterPro; IPR00152; Asx\_hydroxyl.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR002383; GLA\_Blood.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR InterPro; IPR000294; VitK\_dep\_GLA.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF00594; gla; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00001; GLABLOOD.  
 DR SMART; SM00181; EGF; 2.  
 DR SMART; SM00001; EGF\_Like; 2.  
 DR SMART; SM00069; GLA; 1.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE; PS01186; EGF\_2; 2.  
 DR PROSITE; PS01187; EGF\_Ca; 1.  
 DR PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE; PS00240; TRYPsin\_DOM; 1.  
 DR PROSITE; PS00134; TRYPsin\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPsin\_SER; 1.  
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase;  
 KW Hydroxylation; Repeat; Serine protease.  
 SQ SEQUENCE 460 AA; 51784 MW; 0293BC25E9D3ED16 CRC64;

## Alignment Scores:

Pred. No.: 4, 24e-110 Length: 460  
 Score: 1625.00 Matches: 289  
 Percent Similarity: 83.50% Conservativity: 55  
 Best Local Similarity: 70.15% Mismatches: 64  
 Query Match: 69.53% Indels: 4  
 DB: 11 Gaps: 3

US-09-763-153-1 (1-1245) x Q99PC6 (1-460)

QY 1 GCCAACTCCTCTCGAGAGCTCGTCACAGCCCTGGAGCGGAGTGTCATAGAGAG 60  
 42 AlaAsnSerPheLeuGluMetArgProGlySerLeuGluArgGluCysMetGluGlu 61  
 61 ATCTGTGACTTCAGAGCGCCAGGAATTTCCAAATGTGGATGACACACTGGCCTTC 120  
 62 IleCysAspLeuGluGluAlaGlnGluPheGlnAsnValGluAspThrLeuAlaPhe 81  
 QY 121 TGGTCCACAGCCTCGACGGTGACCACTGGTGTGGTCTTGGCCCTGGAGCACCCGTCGCC 180  
 82 TrpIleLysTyrPheAspGlyAspGlnCysSerAlaProLeuAspHisGlnCysAsp 101  
 QY 181 AGCCTGTGTCGGCAGCGGACGACGTGTCATCGACGCGATCGGCAGCTTCAGCTGGCAGTCG 240  
 102 SerProCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysSerCys 121  
 QY 241 CCACAGCGCTGGAGCGGCGCTTCGCCAGCGGAGGTGAGCTTCCTCAATTGCTCGCTG 300  
 122 AspLysGlyTrpGluGlyLysPheCysGlnGlnGluLeuArgPheGlnAspCysArgVal 141  
 QY 301 GACACAGCGGCTGACCGCATTCCTAGAGAGGTGGGTGGCGCGCTGTAGCTGT 360  
 142 AsnAsnGlyGlyCysLeuHisTyrCysLeuGluSerAsnGlyArgArgCysAlaCys 161

QY 361 GGCCTGTCTACAAGCTGGGAGCAGCAGCTCTCTGAGTGTACCCCGCAGTGAAGTTCCT 420  
 162 AlaProGlyTyrGluLeuAlaAspHisMetArgCysLysSerThrValAsnPhePro 181  
 QY 421 TGTGGAGCCCTGGAGCGGATGAGAGAGCGCAGTCACTGCTGAACGAGACACAGAA 480  
 182 CysGlyLysLeuGlyArgTrpIleGluLysLysArgLysIleLeuLysArgAspThr --- 200  
 QY 481 GACCAAGAGACCAAGTA-----GATCCGCGCTCATGTGGGAAGATGACCAAGGGG 534  
 201 AspLeuGluAspGluLeuGluProAspProArgIleValAsnGlyThrLeuThrLysGln 220  
 QY 535 GGAGACAGCCCTGGCAGCGTGTCTCTGGTCTGACTCAAGAGAAAGTGGCTCGCGGCA 594  
 221 GlyAspSerProTrpGlnAlaIleLeuLeuAspSerLysLysLysLeuAlaCysGlyGly 240  
 QY 595 GTGCTCATCCACCCCTCTGGGTGTGTGAGCGGCGCCACTGATGATGATGATGATGATG 654  
 241 ValLeuIleHisThrSerTrpValLeuThrAlaAlaHisCysValGluGlyThrLysLys 260  
 QY 655 CTCCTTGTGAGCTTGGAGATGATGACCTGCGCGCTGGGAGAGTGGGAGCTGGACCTG 714  
 261 LeuThrValArgLeuGlyGluTrpAspLeuArgArgAspHisTrpGluLeuAspLeu 280  
 QY 715 GACATCAAGGAGGTCTCTCTCCACCCCAACTACAGCAAGACCAACCGACATGACATC 774  
 281 AspIleLysGluIleLeuValHisProAsnTyrThrArgSerSerSerAspAsnAspIle 300  
 QY 775 GCACTGCTGACCTGGCCCGCCGCGCCCTCTCCACACCATAGTGCCTCTGCTC 834  
 301 AlaLeuLeuArgLeuAlaGlnProAlaThrLeuSerLysThrIleValProIleCysLeu 320  
 QY 835 CCGACAGCGGCTTCGACAGCGCGAGCTCAATCAGCGCGCGCAGGAGACCTCTGTCAG 894  
 321 ProAsnGlyLeuAlaGln---GluLeuThrGlnAlaGlyGlnGluThrValThr 339  
 QY 895 GGTGCGGTACACAGCAGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 954  
 340 GlyTrpGlyTyrGlnSerAspArgIleLysAspGlyArgArgAsnArgThrPheIleLeu 359  
 QY 955 AACTTCATCAAGATCCCGTGTCCCGCACCAATGAGTGCAGCGAGCTCATGACCAATG 1014  
 360 ThrPheIleArgIleProLeuValAlaArgAsnGluCysValGluValMetLysAsnVal 379  
 QY 1015 GTGTCTGAGAACATGTGTGTGCGGGCATCTCTCGGAGCGGCGGAGGAGGAGGAGG 1074  
 380 ValSerGluAsnMetLeuCysAlaGlyIleIleGlyAspThrArgAspAlaCysAspGly 399  
 QY 1075 GACAGTGGGCGGCGGCTGCTCTCTCCACGCGCAGCTGTTCTCTGTTGGGCTGGTG 1134  
 400 AspSerGlyGlyProMetValValPhePheArgGlyThrTrpPheLeuValGlyLeuVal 419  
 QY 1135 AGCTGGGTGAGGCTGTGGCTCTTCAACACTAGCGGCTTTTACACCAAGTACCGCCG 1194  
 420 SerTrpGlyGlyCysGlyHisThrAsnAsnTyrGlyIleThrThrLysValGlySer 439  
 QY 1195 TACCTGACACTGGATCCATGGGACATCAGAGACAAG 1230  
 440 TyrLeuLysTrpIleHisSerTyrIleGlyGluLys 451  
 RESULT 4  
 Q63207 PRELIMINARY; PRT; 482 AA.  
 ID Q63207  
 AC Q63207  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Factor X.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Theria; Rodentia; Sciurognathi; Muridae; Rattus.  
 NCBI\_TaxID=10116;



088947 ID O88947 PRELIMINARY; PRT; 481 AA.

AC O88947;

DT 01-NOV-1998 (TRENBLRel. 08, Created)

DT 01-NOV-1998 (TRENBLRel. 08, Last sequence update)

DT 01-MAR-2002 (TRENBLRel. 20, Last annotation update)

DE Coagulation factor X precursor.

GN F10.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI\_TaxID=10090;

ON [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL6 X CBA; TISSUE=LIVER;

RX MEDLINE=98347933; PubMed=9684791;

RA Liang Z., Cooper A., DeFord M.E., Carmeliet P., Collen D.,

RA Castellino F.J., Rosen E.D.

RT "Cloning and characterization of a cDNA encoding murine coagulation

RT factor X";

RT Thromb. Haemost. 80:87-91(1998).

CC [2]

CC SEQUENCE FROM N.A.

CC STRAIN=129Sj;

CC Cooper A., Liang Z., Castellino F.J., Rosen E.D.;

CC "Cloning and Characterization of the Murine Factor X Gene.";

CC Thromb. Haemost. 0:0-0(2000).

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

CC TRYPSIN FAMILY.

DR EMBL; AF087644; AAC36345.1; -.

DR EMBL; AF211347; AAF22980.1; -.

DR HSSP; P00742; 1XKA.

DR MEROPS; S01.216; -.

DR MGD; MGI:103107; F10.

DR InterPro; IPR000152; Asx\_hydroxyl.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR000742; EGF\_2.

DR InterPro; IPR001881; EGF\_Ca.

DR InterPro; IPR002383; GLA\_blood.

DR InterPro; IPR001254; Ser\_protease\_Try.

DR InterPro; IPR000294; VitK\_dep\_GLA.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00594; gla; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00001; GLABLOOD.

DR SMART; SM00179; EGF\_Ca; 1.

DR SMART; SM00001; EGF\_like; 1.

DR SMART; SM00089; GLA; 1.

DR SMART; SM00020; Tryp\_Spc; 1.

DR PROSITE; PS00010; ASX\_HYDROXYL; UNKNOWN\_1.

DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.

DR PROSITE; PS01186; EGF\_2; 2.

DR PROSITE; PS01187; EGF\_Ca; 1.

DR PROSITE; PS00011; GLU\_CARBOXYLATION; 1.

DR PROSITE; PS02040; TRYPSIN\_DOM; 1.

DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.

DR PROSITE; PS00135; TRYPSIN\_SER; 1.

KW Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Repeat;

KW Serine protease; Signal.

FT SIGNAL 1 40

FT CHAIN 41 481

SQ SEQUENCE 481 AA; 54018 MW; 8AC09DE5EF9D271E CRC64;

Alignment Scores:

Score: 1.31e-49 Length: 481

Matches: 794.00

Percent Similarity: 53.71% Conservative: 77

Best Local Similarity: 36.40% Mismatches: 158

Query Match: 33.98% Indels: 48

DB: 11 Gaps: 9

US-09-763-153-1 (1-1245) x 088947 (1-481)

QY 1 GCCAACTCTCTCTGGAGGAGCTCCGTCACAGCAGCCTGGAGCGGAGTGCATAGAGGAG 60

DB 41 AlaAsnSerPheGluGluPheLysGlyAsnLeuGluGluGluCysMetGluGlu 60

QY 61 ATCTGTGACTTGGAGGAGCCAAAGAAATTTCCAAAATGTGGATGACATGGCGCTTC 120

DB 61 IleCysSerTyrGluGluValArgGluIlePheGluAspGluLysThrLysGluTyr 80

QY 121 TGTCTCAAGCAGCTCGACGGTGACCATGCTGTGCTTCTGCCCTTGGAGCACCOCGTCGCC 180

DB 81 TrpThrLysTyrLysAspGlyAspGlnCysGluSerPro----- 94

QY 181 AGCTGTGCTGCGGCGGACGCGCAGCTGCATCGAGCGATCGAGCTTCAGTCCGACTGC 240

DB 95 -----CysGlnAsnGlnGlyAlaCysA-ArgPheGlyIleGlyTyrThrCysThrCys 112

QY 241 CGCAGCGGCTGGAGGCGGCTTCTGCCAGCGGAGGTGAGTTCCTCAATGCTCGCTG 300

DB 113 SerGluGlyPheGluGlyLysAsnCysGluLeuPheValArgLysLeu--CysArgLeu 131

QY 301 GACAACGCGGCTGCACGCTTACTGCTAGAGGAGTGGGCGGCGCTGTAGCTGT 360

DB 132 AspAsnGlyAspCysAspGlnPheCysArgGluGluGlnAsnSerValValCysSerCys 151

QY 361 GCGCTGGCTTACAGCTGGGGGACGACCTCTCCACAGTGTCAACCCGCGAGTGAAGTTCCT 420

DB 152 AlaSerGlyTyrPheLeuGlyAsnAspGlyLysSerCysIleSerThrAlaProPhePro 171

QY 421 TGTGGGAGCGCTGGAGCGGATGGAGAGAAG-----CGCATCTACCTG 465

DB 172 CysGlyLysIleThrThrGlyArgArgLysArgSerValAlaLeuAsnThrSerAspSer 191

QY 466 AAACGAGACACAAACAC-----CAAGAGACCAAGTAGATCCG----- 504

DB 192 GluLeuAspLeuGluAspAlaLeuLeuAspGluAspPheLeuSerProThrGluAsnPro 211

QY 505 -----CGG 507

DB 212 IleGluLeuLeuAsnLeuAsnGluThrGlnProGluArgSerSerAspAspLeuValArg 231

QY 508 CTCATTGATGGGAAGATGACACGCGGGGAGACAGCCCTGGCAGGTGTCTGTCTGGAC 567

DB 232 IleValGlyGlyArgGluCysLysAspGlyGluCysProTrpGlnAlaLeuLeuLeuLeu 251

QY 568 TCAAAGAGAAGCTGGCTGCGGGGCGAGTCTCATCCACCCCTCTGCTGGTGTGACAGG 627

DB 252 GluAspAsnGluGlyPheCysGlyGlyThrIleLeuAsnGluPheTyrIleLeuThrAla 271

QY 628 GCCCACTGATGATGATCCAAAGAGCTCTGTGTCAGGCTGGAGAGTATGACCTGCGG 687

DB 272 AlaHisCysLeuHisGlnAlaArgArgPheLysValArgValGlyAspArgAsnThrGlu 291

QY 688 CGTGGGAGAGTGGGAGCTGGACCTGGACATCAAGAGGTCTTCCTCCACCCCAACTAC 747

DB 292 LysGluGluGlyAsnGluMetValHisGluValAspValValIleLysHisAsnLysPhe 311

QY 748 AGCAAGAGCACCACCAATGATCGCAGTCTCTCCACCTGGCCCGCCGCCACCCCTC 807

DB 312 GlnArgAspThrTyrAspTyrAspIleAlaValLeuArgLeuLysThrProIleThrPhe 331

QY 808 TCGCAGACCATAGTGCCTATCTGCTCCCGGACGAGCGGCTTGCAGAGCGGAGTCAAT 867

DB 332 ArgMetAsnValAlaProAlaCysLeuProGlnLysAspTrpAlaGluSerThrLeu--- 350

QY 868 CAGCGCGGCGAGGAGACC---CTCTGTAGCGGCTGGGGTACACAGCAGCCGAGAGAG 924

DB 351 ---MetThrGlnLysThrGlyIleValSerGlyPheGly-----ArgThrHisGluLys 367

QY 925 GAGGCCAAGAGAACCACCGACCTTCGTCCTCAACTTCATCAAGATTCCTGGTCCCGCAC 984

DB 368 GlyArgGlnSerAsn-----IleLeuLysMetLeuGluValProTyrValAspArg 384











Db 321 LeuCysLeuProGluArgThrPheSerGluArgThrLeuAlaPheVal----ArgPheSer 339  
 QY 886 CTCGTGACGGCTGGGGCTTACCAGCAGCAGCCGAGAGAGAGGCCAAGAACCGCACCC 945  
 |||||:|||||  
 Db 340 LeuValSerGlyTrpGlyGlnLeuLeuAspArgGlyAlaThrAla----- 354  
 QY 946 TTGCTCTCAACTTCATCAAGATTCCCGTGGTCCCGCACAATGAGTGCAGCGAGTCAATG 1005  
 |||||:|||||  
 Db 355 LeuGluLeuMetValLeuAsnValProArgLeuMetThrGlnAspCysLeuGlnGlnSer 374  
 QY 1006 -----AGCAACATGTTCTGTGAGAACATGCTGTCTCGGGCATCTCTCGGG 1050  
 |||||:|||||  
 Db 375 ArgLysValGlyAspSerProAsnIleThrGluTyrMetPheCysAlaGlyTyrSerAsp 394  
 QY 1051 GACCGCAGATGCTCGCAGGCGACAGTGGGGGCCCATGGTGCCTCTTCCACGGC 1110  
 :|||||:|||||  
 QY 395 GlySerLysAspSerCysAlaGlyAspSerGlyGlyProHisAlaThrHisTyrArgGly 414  
 QY 1111 ACTGTGTTCTGCTGGGCTGTGAGTGGGTGGGTGGGTGGGTCTCTTCCACAACTAC 1170  
 |||||:|||||  
 Db 415 ThrTrpTyrLeuThrGlyIleValSerTrpGlyGlnGlyCysAlaThrValGlyHisPhe 434  
 QY 1171 GCGTTTACCAAGTACGCGCTACCTGACGTGATCCATGGCCACATCAGACACAAG 1230  
 |||||:|||||  
 Db 435 GlyValTyrThrArgValSerGlnTyrIleGluTrpLeuGlnLysLeuMetArgSerGlu 454  
 QY 1231 GAAGCCCC 1239  
 |||||  
 Db 455 ProArgPro 457  
 RESULT 10  
 Q9GMD9 PRELIMINARY; PRT; 469 AA.  
 ID Q9GMD9  
 AC Q9GMD9  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Coagulation factor X.  
 OS Ornithorhynchus anatinus (Duckbill platypus).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.  
 OX NCBI\_TaxID=9258;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21015017; PubMed=11132153;  
 RA Poorafshar M., Aveskogh M., Munday B., Hellman L.;  
 BA "Identification and structural analysis of four serine proteases in a  
 monotrema, the platypus, Ornithorhynchus anatinus.";  
 JL Immunogenetics 52:19-28(2000).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 TRYPSIN FAMILY.  
 CC EMBL: AF275654; XAG00453.1; -.  
 DR HSPF; P00742; 1XKB.  
 DR MEROPS; S01.216; -.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR000742; EGF-2.  
 DR InterPro; IPR001881; EGF-Ca.  
 DR InterPro; IPR002383; GLA\_blood.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF00594; gla; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00001; GLABLOOD.  
 DR SMART; SM00181; EGF; 2.  
 DR SMART; SM00179; EGF\_CA; 1.  
 DR SMART; SM00001; EGF-like; 2.  
 DR SMART; SM00069; GLA; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; UNKNOWN\_1.

DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE; PS01186; EGF\_2; 2.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 DR PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 SQ Hydrolase; Serine protease.  
 SQ SEQUENCE 469 AA; 52196 MW; 4C66C230D0758F6A CRC64;  
 Alignment Scores:  
 Pred. No.: 4,39e-48 Length: 469  
 Score: 773.00 Matches: 162  
 Percent Similarity: 53.27% Conservative: 74  
 Best Local Similarity: 36.57% Mismatches: 157  
 Query Match: 33.08% Indels: 50  
 DB: Gaps: 10  
 US-09-763-153-1 (1-1245) x Q9GMD9 (1-469)  
 QY 1 GCCAACTCTTCCTGGAGGAGTCCGTCCACAGCAGCCTGGAGCGGAGTCATAGAGGAG 60  
 |||||  
 Db 41 AlaAsnSerLeuPheGluGluLeuLysLysGlyAsnLeuGluArgGluCysAsnGluGlu 60  
 QY 61 ATCTGTGACTTCGAGGAGGCCAAGGAAATTTCCAAATGTGGATGACACACTGGCCTTC 120  
 |||||  
 Db 61 ThrCysSerTyrGluGluAlaArgGluValPheGluAspThrAspLysThrAsnGluPhe 80  
 QY 121 TGTCTCAAGCAGCTCGACGCTGACCTGCTTGTGCTTGTGCTTGTGGAGCACCCTGGGCC 180  
 |||||  
 Db 81 TrpAsnIleTyrLysAspGlyAsnGlnCysGluThrGlnPro----- 94  
 QY 181 AGCCTGTGTCGGGGCAGCGACGCTGCATCAGCAGCATCGCAGCTTCAGCTTCGACTGC 240  
 |||||  
 Db 95 -----CysGlnAsnGlnGlyValCysLysAspGlyLeuAlaGluTyrThrCysLeuCys 112  
 QY 241 CGCAGCGCTGGGAGCGCTTCTGCCAGCGGAGGTGAGCTTCCTCAATTGCTCGCTG 300  
 |||||  
 Db 113 SerAlaGlyTyrGluGlyLysAsnCysAspThrThrThrValLysIle---CysSerLeu 131  
 QY 301 GACAACGCGCTGCACGCATTACTGCCTAGAGAGGTGGCTGGCGCGCTGTAGCTGT 360  
 |||||  
 Db 132 AsnAsnGlyAspCysGluGlnPheCysLysSerValValAsnThrValValCysSerCys 151  
 QY 361 GCGCTCGCTACAGCTGGGAGGAGCAGCTCCTCAGCTGTCACCCGCGAGTGAATTCCT 420  
 |||||  
 Db 152 AlaGlnGlyTyrIleLeuGlyAspGlnLysSerCysIleProThrValProPhePro 171  
 QY 421 TGTGGAGGCGCTCGAAGCGATGGAGAAGAAG----- 453  
 |||||  
 Db 172 CysGlyLysLeuThrValGlyArgArgLysArgSerArgGluLeuProGluGluGlnAsp 191  
 QY 454 -----CCAGTCACTGAAACAGAGAC-----ACAGAAGAACCAAGAACCA 495  
 |||||  
 Db 192 GlyAspAsnAlaHisValAlaGluAspValLeuGluAlaThrGluAsnProPheGluAla 211  
 QY 496 GTAGATCCG-----CGGCTC 510  
 |||||  
 Db 212 GluProProAspGlnAsnThrThrLeuAlaGluProGlyGluAsnAlaLeuValArgIle 231  
 QY 511 ATTGATGGGAAGATGACACGCGGGGAGACAGCCCTCGGAGGTGGTCTCTGCTGACAGCG 567  
 |||||  
 Db 232 ValGlyArgGluCysHisAspGlyGluCysProTrpGlnAlaLeuValAsnAsp 251  
 QY 568 TCRAAGAAGAAGCTGGCTCGCGGGCAGTGTCTATCCACCCCTCGGCTGGTCTGACAGCG 627  
 |||||  
 Db 252 GluAsnGlyGlnGlyPheCysGlyGlyThrIleLeuAsnGluTyrIleLeuSerAla 271  
 QY 628 GCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 697  
 |||||  
 Db 272 AlaHisCysMethHisGlnAlaLysArgPheLysValArgValGlyGluArgAspThrGlu 291  
 QY 688 CGCTGGGAGAAGTGGGAGCTGGACCTGGACATCAAGAGGTCTTCGTCCACCCCACTAC 747

```
Db 292 LysAspSerSerGluMetAlaHisGluValGluLysValIleValHisSerLysPhe 311
    ::: ::: |||::: ::::::::::: ||| |||:::
Qy 748 AGCAGAGCACCACCGACATGACGCTGCTGACCTGGCCCGCCGACCCCTC 807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 312 ValLysThrThrAspPheAspIleAlaValIleLysLeuLysThrProIleThrPhe 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 808 TCGCAGACCATAGTGCCTCCTCCGACAGCGCGCTGTCAGAGCGCGAGCTC--- 864
    ::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 332 ArgMetAsnValSerProAlaCysLeuProGluLysAspTrpAlaGluAspIleLeuMet 351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 865 ---AATCAGCGCGCGAGGACCCCTGTCAGCGGCTGGGCTACCCAGCAGCGCGAGAG 921
    ::||| ||| ::||| ||| ::||| ||| ::||| ||| ::||| |||
Db 352 AsnGlnLysAlaGly-----ValValSerGlyPheGly-----ArgVal 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 922 AAGGAGGCGCAAGAACCGCACCTGCTCCTCAACTTCATCAGATTCCTGGTCCCG 981
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 365 HisGluLysGlyArgProSerThr---ValLeuLysMetLeuGluValProTyrValGlu 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 982 CACAATGAGTCAGCGAGGTCATGAGCAACATGCTGTGAGAACATGCTGTGCGGGC 1041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 384 ArgThrThrCysLysGlnSerSerSerPheAspIleThrProAsnMetPheCysAlaGly 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1042 ATCTCGGGGACCGCAGGATGCTGCGAGGCGCACAGTGGGGGCCCATGGTCCCTCC 1101
    ::||| ||| ::||| ||| ::||| ||| ::||| ||| ::||| |||
Db 404 TyrAspSerArgProGluAspAlaCysGlnGlyAspSerGlyGlyProHisValThrLys 423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1102 TTCACGCGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424 TyrLysAspThrThrPheValThrGlyIleValSerTrpGlyGluGlyCysAlaGlnAsn 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1162 CACAACACTACGCGTTACACAAAGTCAGCGCTACCTCGAGCTGATCCATGGGCACATC 1221
    ::||| ||| ::||| ||| ::||| ||| ::||| ||| ::||| |||
Db 444 GlyLysPheGlyValTyrThrLysAlaAlaThrPheLeuSerTrpIleLysArgMetMet 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1222 AGAGCAAG 1230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 464 ArgGlnLys 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
Q61109 PRELIMINARY; PRT; 446 AA.
ID Q61109
AC Q61109;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Coagulation factor VII.
GN F7 OR FVII.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
TISSUE=LIVER.
MEDLINE=96276538; PubMed=8701412;
RA Idusogie E., Rosen E., Geng J.P., Carmeliet P., Collen D.,
RA Castellino F.J.;
RT Thromb. Haemost. 75:481-487(1996).
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; U44795; AAC52570.1; -.
DR HSSP; P08709; IFAK.
DR MEROPS; S01.215; -.
DR MGD; MGI:109325; F7.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR002383; GLA_blood.
```

```
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; Vitk_dep_GLA.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00594; gla; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0001; GLABLOOD.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00001; EGF_Like; 1.
DR SMART; SM00069; GLA; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00011; GUL_CARBOXYLATION; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Repeat;
KW Serine protease.
SQ SEQUENCE 446 AA; 50318 MW; 482FD09BEFDA6870 CRC64;

Alignment Scores:
Pred. No.: 5,15e-48 Length: 446
Score: 772.00 Matches: 163
Percent Similarity: 56.43% Conservative: 74
Best Local Similarity: 38.81% Mismatches: 147
Query Match: 33.03% Indels: 36
DB: 11 Gaps: 11

US-09-763-153-1 (1-1245) x Q61109 (1-446)
Qy 1 GCACACTCTCTCTGGAGGAGCTCCGTCACAGCAGCTGGAGCGAGTCATAGAGGAG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 42 AlaAsnSerLeuLeuGluLeuTrpProGlySerLeuGluArgGluCysAsnGluGlu 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 61 ATCTGTGACTTCGAGGAGGCCAAGGAAATTTCCAAATGTGGATGACACATGCGCTTC 120
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 GlnCysSerPheGluGluAlaArgGluIlePheLysSerProGluArgThrLysGlnPhe 81
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 121 TGTCTCAACAGCAGCTCGACGCTGACCATCTTGGTCTTCCCTTGGAGCAGCCGCGCC 180
    ||| ::| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 82 TrpIleValTyrSerAspGlyAspGlnCys-----AlaSerAsnProCysGln 97
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 AGCCTGTGCTGCGGCGCACGCGCAGTCATCGCGGATCGGAGCTTCAGCTCGCAGCTC 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 98 AsnVal-----GlyThrCysGlnAspHisLeuLysSerTyrValCysPheCys 113
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 241 CGCAGCGCTGGAGGCGCGCTTCTCCAGCGGAGGTGAGC---TTCCTCAATTGCTCG 297
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 114 LeuLeuAspPheGluGlyArgAsnCysGluLysSerLysAsnGluLeuIleCysAla 133
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 298 CTGACAGCGCGGCTGCGACGATTAATCTAGAGAGGCTGGCTGGCGGCGC---TGT 354
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 134 AsnGluAsnGlyAspCysAspGlnTyrCysArgAspHisValGlyThrLysArgThrCys 153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 355 AGCTGTGCGCCTGCCTCAAGCTGGGCGGACGACCTCTCTGACGTGTCACCCCGACGTGAAG 414
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 154 SerCysHisGluAspTyrThrLeuGlnProAspGluValSerCysLysProLysValGlu 173
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 415 TTCCTTGTGGGAGGCGCTGGAGAGCGGATGGAGAGACGCGGATCCTCCTGAACAGAC 474
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 174 TyrProCysGlyArg---IleProValValGluLysArgAsnSerSerArgGlnGly 192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 475 ACAGAGACCAAGACAGCAAGTAGATCCGCGCTCATTTGATGGGAGAGATGACCGGCGG 534
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 193 -----ArgIleValGlyCysAsnValCysProLys 202
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 535 GGAGACAGCCCTGGCAGGTGGTCTGCTGAGTCAAGAGAGAGGCTGCGCTCGCGGGGA 594
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 203 GlyGluCysProTrpGlnAlaVal---LeuLysIleAsnGlyLeuLeuLysGlyAla 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```





```
QY 472 -----CACACAGAAGACCAAGAACACCAAGTAGAT----- 501
Db 196 TyrValAsnSerThrGluAlaGluThrIleLeuAsnIleThrGlnSerThrGlnSer 215
QY 502 -----CCGGCGCTCATTTGATGGGAAGATGACACAGGGGGGAGACAGCCCTGG 549
Db 216 PheAsnAspPheThrArgValValGlyGlyGluAspAlaLysProGlyGlnPheProTyr 235
QY 550 CAGGTGCTCTGCTGGACTCAAGAGAAGAGCTGGCCCTGGGGCAGTGTCTATCCACCC 609
Db 236 GlnValVal-----LeuAsnGlyLysValAspAlaPheCysGlyGlySerIleValAsnGlu 254
QY 610 TCCTGGTGTCTGACAGCGGCCACTGCTGATGAGTCAAGAGAGCTCTTGTTCAGGCTT 669
Db 255 LysTrpIleValThrAlaAlaHisCysValGluThrGlyValIleThrValAla 274
QY 670 GGAGATGATGACCTGGCGGCTGGGAGAGTGGGAGCTGGACCTGGACATCAAGAGGTC 729
Db 275 GlyGluHisAsnIleGluGluThrGluHisThrGluGlnLysArgAsnValIleArgIle 294
QY 730 TTCGTCCACCCCACTACACAGACACACACC-----GACAATGACATCCGACTGTG 783
Db 295 IleProHisHisAsnTyrAsnAlaAlaIleAsnLysTyrAsnHisAspIleAlaLeu 314
QY 784 CACCTGGCCCGAGCCGCCCTCTCGACACCATAGTCCCATCTGCTCCCGGACACG 843
Db 315 GluLeuAspGluProLeuValLeuAsnSerTyrValThrProIleCysIleAlaAspLys 334
QY 844 GGCTTGGCAGCGCGAGCTCAATCAGCGCGCCAGGAGACCTCGTGACGGGCTGGGG 903
Db 335 GluTyrThrAsnIlePheLeuLysPheGly-----SerGlyTyrValSerGlyTrpGly 352
QY 904 -----TACCACAGCAGCGAGAGAGAGGAGCCACAGAAACCGACCTTCTCTCAAC 957
Db 353 ArgValPheHisLysGlyArgSer-----AlaLeuValLeuGln 365
QY 958 TTCATCAAGATTCCCGTGTCCCGCAATGAGTGCAGCGGAGTGCATGAGCAACATGGT 1017
Db 366 TyrLeuArgValProLeuValAspArgAlaThrCysLeuArgSerThrLysPheThrIle 385
QY 1018 TCTGAGAATGCTGTGTGGGGCATCTCCGGGAGCCGAGAGTCCCTCGCAGGGCGAC 1077
Db 386 TyrAsnAsnMetPheCysAlaGlyPheHisGluGlyGlyArgAspSerCysGlnGlyAsp 405
QY 1078 AGTGGGGGCCCATGCTGCTCCCTCTCCACGCGACCTGTTCTGCTGGTGGGCTGTGAGC 1137
Db 406 SerGlyGlyProHisValThrGluValGluGlyThrSerPheLeuThrGlyIleSer 425
QY 1138 TGGGTGAGGGGCTGTGGGCTCTTCACTACAGCGGTTTACACAAAGTACAGCGCTAC 1197
Db 426 TrpGlyGluGluCysAlaMetLysGlyLysTyrGlyIleTyrThrLysValSerArgTyr 445
QY 1198 CTCGACTGATC 1209
Db 446 ValAsnTrpIle 449
RESULT 14
Q95ND6 PRELIMINARY; PRT; 461 AA.
AC Q95ND6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Coagulation factor XI.
GN F9.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=505;
```

```
RA Satta Y.;
RT "Comparison of DNA and protein polymorphisms between humans and
RL chimpanzees.";
DR EMBL; AB062471; BAB58886.1; .
DR EMBL; AB062459; BAB58886.1; JOINED.
DR EMBL; AB062461; BAB58886.1; JOINED.
DR EMBL; AB062463; BAB58886.1; JOINED.
DR EMBL; AB062465; BAB58886.1; JOINED.
DR EMBL; AB062467; BAB58886.1; JOINED.
DR EMBL; AB062469; BAB58886.1; JOINED.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00594; gla; 1.
DR Pfam; PF00089; trypsin; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_2.
DR PROSITE; PS01187; EGF_CA; UNKNOWN_1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; UNKNOWN_1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
SQ Hydrolase; Serine protease.
KW SEQUENCE 461 AA; 51695 MW; 8F5A69A525DF65B5 CRC64;

Alignment Scores:
Pred. No.: 4.41e-44 Length: 461
Score: 718.00 Matches: 149
Percent Similarity: 50.94% Conservative: 67
Best Local Similarity: 35.14% Mismatches: 162
Query Match: 30.72% Indels: 46
DB: 6 Gaps: 10

US-09-763-153-1 (1-1245) x 095ND6 (1-461)
QY 13 CTGAGGAGCTCCCTCACAGCAGCTGGAGGGAGTGTCATAGAGAGATCTGTGACTTC 72
Db 52 LeuGluGluPheValGlnGlyAsnLeuGluArgGluCysMetGluGluLysCysSerPhe 71
QY 73 GAGGAGGCCAAGAAATTTCCAAATGTGGATGACACTGGCTTCTGTGTCCAAAGCAC 132
Db 72 GluGluAlaArgGluValPheGluAsnThrGluArgThrThrGluPheTrpLysGluTyr 91
QY 133 GTCGACGGTGACCACTGCTTGGTCTTGCCTTGGACACCGCTGGCCGACCTGTGCTGC 192
Db 92 ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu 103
QY 193 GGGCAGCGCAGCTGATCGACGCGCTTACGCTGCGACTGCGCAGCGGCTGG 252
Db 104 AsnGlyGlySerCysLysAspAspIleAsnSerTyrGluCysTrpCysProPheGlyPhe 123
QY 253 GAGGGCGCTTCTCCAGCGGAGGTAGCTTCTCAATGCTGCTGGAGCAAGCGGCG 312
Db 124 GluGlyLysAsnCysGluLeuAspValThr-----CysAsnIleLysAsnGlyArg 140
QY 313 TGCACGCTACTTCTAGAGGAGGTGGCTGGCGG---CGCTAGCTGCTGGCCCTGGC 369
Db 141 CysGluGlnPheCysLysAsnSerAlaAspAsnLysValValCysSerCysThrGluGly 160
QY 370 TACAAGCTGGGGGACGACCTCTCTGCTGACCTCCAGCTGACAGTTCCTTGTGGGAGG 429
Db 161 TyrArgLeuAlaGluAsnGlnLysSerCysGluProAlaValProPheProCysGlyArg 180
QY 430 CCCTGGAGCGGATGGAGAGACGACGCTACCTGAAA----- 468
Db 181 ValSerValSerGlnThrSerLysLeuThrArgAlaGluThrValPheProAspValAsp 200
```

```

QY 469 -----CGAGACACAGAA-----GACCAAGAA 489
Db 201 TyrValAsnSerThrGluAlaGluThrIleLeuAspAsnIleThrGlnThrProGlnSer 220
QY 490 GACCAAGTAGATCCCGGCTCATTCATGGAGATGACCGGGGGGAGACAGCCCTGG 549
Db 221 *****ThrArg***ValGlyGluAsp***AsnProGlyGlnTyrPro--- 239
QY 550 CAGGTGGCTCTGCTGACTCAAGAGAGCTGGCTCGGGGAGTGTCTCATCACCCC 609
Db 240 Arg***ValValLeuAsnGlyLysValAspAlaPheCysGlySerIleValAsnGlu 259
QY 610 TCTGGTGGTCTGACAGCGCCACTGTCATGATGAGTCCAAAGAGCTCTGTGAGGCT 669
Db 260 LysTrpIleValAlaAlaHisCysValAspThrGlyValLysIleThrValValAla 279
QY 670 GGAGAGTAGACTCGCGGCTGGGAGAGTGGAGCTGGACCTGGACATCAAGAGGTC 729
Db 280 GlyGluHisAsnIleGluThrGluHisThrGluGlnLysArgAsnValIleArgIle 299
QY 730 TTCGTCCACCCCACTACAGCAAGAGACACC-----GACAATGACATCGCATGCTG 783
Db 300 IleProHisHisAsnTyrAsnAlaIleAsnLysTyrAsnHisAspIleAlaLeuLeu 319
QY 784 CACCTGGCCACCCGCCACCTCTCGCAGACCATAGTCCCATCTGCTCCCGGACAGC 843
Db 320 GluLeuAspGluProLeuValLeuAsnSerTyrValThrProIleCysIleAlaAsp 339
QY 844 GGCCTGACAGCGGCTCAATCAGCGCGCCAGGAGACCTCGTCAAGGCTGGGGC 903
Db 340 GluTyrThrAsnIlePheLeuLysPheGly-----SerglyTyrValSerGlyTrpGly 357
QY 904 -----TACCACGACGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 957
Db 358 ArgValPheHisGlyArgSer-----AlaLeuValLeuGln 370
QY 958 TTCATCAAGATCCCGTGGTCCCGCACAATGATGTCACGAGGTCATGAGCAACATG 1017
Db 371 TyrLeuArgValProLeuValAspArgAlaThrCysLeuArgSerThrLysPheThrIle 390
QY 1018 TCTGAGACATCTGTGGGCGGATCTCGGGGACCGGAGGAGGAGGAGGAGGAGG 1077
Db 391 TyrAsnAsnMetPheCysAlaGlyPheHisGluGlyArgAspSerCysGlnGlyAsp 410
QY 1078 ATGGGGGCGCCATGCTGCTCCCTCCCTCCACGCGACCTGCTGCTGGTGGTGGTGG 1137
Db 411 SerGlyGlyProHisValThrGluValGluGlyThrSerPheLeuThrGlyIleIleSer 430
QY 1138 TGGGGTGAGGCTGTGGGCTCCTTCACACTACGGCGTTTACACCAAGTACGCGCTAC 1197
Db 431 TrpGlyGluGluCysAlaMetLysGlyLysTyrGlyIleTyrThrLysValSerArgTyr 450
QY 1198 CTCGACTGGATC 1209
Db 451 ValAsnTrpIle 454

RESULT 15
Q9PTW7 PRELIMINARY; PRT; 608 AA.
AC Q9PTW7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Prothrombin.
GN OSPT.
OS Struthio camelus (Ostrich).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Palaeognathae; Struthioniformes; Struthionidae;
OC Struthio.
OX NCBI_TaxID=8801;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;

```

```

RX MEDLINE=20579470; PubMed=11137455;
RA Frost C., Naude R., Oelofsen W., Muramoto K., Naganuma T., Ogawa T.;
RT "Purification and characterization of ostrich prothrombin.";
RL Int. J. Biochem. Cell Biol. 32:1151-1159(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB028871; BAA89046.1; -.
DR HSSP; P00734; IUVS.
DR MEROPS; S01.217; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002383; GLA_blood.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00594; gla; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00001; GLABLOOD.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00069; GLA; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; tryp_spc; 1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 608 AA; 69392 MW; 11B974B9AE54EA2 CRC64;

Alignment Scores:
Pred. No.: 5,15e-30 Length: 608
Score: 525.00 Matches: 139
Percent Similarity: 39.00% Conservative: 59
Best Local Similarity: 28.44% Mismatches: 175
Query Match: 22.46% Indels: 166
DB: 13 Gaps: 19

US-09-763-153-1 (1-1245) x Q9PTW7 (1-608)
QY 10 TTCCTGAGGAGTCCCTCAGCAGCGCTCGAGCGGAGTGCATAGAGGAGATCTGTGAC 69
Db 48 PheLeuGluMetLeuLysGlyAsnLeuGluArgGluCysLeuGluIleCysIle 67
QY 70 TTCGAGAGGCCAAGGAATTTCCAAATGTGGATCACACACTGGCCTTCTGTCTCAAG 129
Db 68 TyrGluAlaPheGluAlaLeuGluSerThrAlaArgThrGluGluPheTrpSerLys 87
QY 130 CAC----- 132
Db 88 TyrGlnAlaCysGlnGlyIleArgLysSerArgThrValLeuAspAlaCysLeuGluGly 107
QY 133 -----GTCGACGGT 141
Db 108 AsnCysSerThrAspLeuGlyGlnAsnTyrArgGlyThrIleSerHisThrLysSerGly 127
QY 142 GACCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 174
Db 128 ThrGluCysGlnMetTrpThrSerLysTyrProHisIleProLysPheAsnThrThrIle 147
QY 174 ----- 174
Db 148 HisProAsnLeuIleGluAsnTyrCysArgAsnProAspAsnProGluGlyProTrp 167
QY 175 -----TGCGCCAGCCCTGCTCGCGGAC 198
Db 168 CysTyrThrArgAspProThrValProArgGluCysProIleProValCysGlyGlu 187
QY 198 ----- 198

```



188	Db	GluaRgThrThrValGluPheThrProArgValLysProProAlaSerThrGluProCys	207
199	QY	-----GGCAGCTGCATCGACGCGATCGCGAGCTTCAGC	231
208	Db	GlucInGluLysGlyMetLeuTyrAlaGlyThrLeuSerValThrIleSerGlyAlaLys	227
232	QY	TGC-----GACTGCGCAGCGCTGGGAGGCGCGTCTGCCAGCGC	273
228	Db	CysLeuProTrpAsnSerGluLysAlaLysGluValLeuGlnGlyLysThrIleLeuThr	247
274	QY	GAGTGAGCTTCTCTCAAT-----	291
248	Db	GluValLysLeuLeuGluAsnTyrCysArgAsnProAspAlaAspGluGlyValTrp	267
292	QY	TGCTGCTGGACAAACGGCGCTGCAGCATTAATCTGC-----	327
268	QY	CysValThrAspLeuProHisPheGluTyrCysAspLeuGlnTyrCysAspSerSer	287
328	QY	-----CTAGAGGAGTGGCTGGCGCGCTGTAGCTGTGCGCGCT	366
288	Db	LeuGluAspGlyAsnGluGlnLeuGluIleSerGlyArg-----ThrValLeuGln	305
367	QY	GGCTACAAG--CTGGGGGACGACCTCTGCGAGTGTCAACCGCGAGTCAAGTTCCTTGT	423
306	Db	GluTyrLysThrPhePheAsp-----LysThrPheGlySerGlyAlaAspCys	323
424	QY	GGG--AGGCGCTGGAAGCGGATGGAGAAGAGCGCAGTCACTGAAACGAGACACAGAA	480
324	Db	GlyLeaArgPro-----LeuPheGluLysLysLysLysLysLysSerGlyLysGlu	341
481	QY	GACCAAGAAGACCAATAGATCCGGCGTCAATTGATGGGAAGATGACACGCGGGGAGAC	540
342	Db	LeuLeuGluSerTyrIleGlySerArgValValHisGlyAspAlaGluValGlySer	361
541	QY	AGCCCTGCGCAGGTGCTCTGCTG--GACTCAAGAGAAGCTGGCGCTCGCGGGCAGTG	597
362	Db	SerProTrpGlnValMetLeuTyrLysLysSerProGlnGluLeuCysGlyAlaSer	381
598	QY	CTCATCCACCCCTCTCGGTGCTGACAGCGGCCACTGCATG-----GAT	642
382	Db	LeuIleSerAspSerTrpValLeuThrAlaAlaHisCysLeuTyrTyrProTrpTrpAsp	401
643	QY	GAG-----TCCAAGAAGTCTCTCTAGGCTTGGAGAGTATGACCTGCGCGCGTGG	693
402	Db	LysAsnLeuThrThrAsnAspIleLeuValArgIleGlyLysHisPheArgAlaLysTyr	421
694	QY	GACAAGTGGGAGCTGGACCTGAC--ATCAGGAGGTCTTCGTCCACCCCACTACAGC	750
422	Db	GluLysAsnLysGluLysIleAlaLeuLeuAspLysIleIleHisProLysTyrAsn	441
751	QY	---AAGAGCACCACGACAATGACATCGACTGCTGCACTGGCGCCAGCGCCACCCCTC	807
442	Db	TrpLysGluAsnMetAspArgAspIleAlaLeuMetHisLeuLysArgProValIlePhe	461
808	QY	TGCGAGACCATAGTGCCTATCTCCCTCCGGACACGCGCCTTGCAGACGCGAGCTCAAT	867
462	Db	SerAspTyrIleHisProValCysLeuProThrLysGluValGlnArg--LeuMet	480
868	QY	CAGCGCGCGCAGAGACCTCTGTGAGCGGTGGGCTACCAACACGACCGCAGAGAAGGAG	927
481	Db	LeuAlaGlyTyrLysGlyArgValThrGlyTrpGlyAsnLeuLysGluThrTrpAlaThr	500
928	QY	GCCAAGAGAAACGGCACCTTCGTCTCAACTTCATCAAGATTCCGCTGGTCCCGACAAT	987
501	Db	AsnProSerAsnLeuProThrValLeuGlnLeuAsnValProIleValAspGlnAsp	520
988	QY	GAGTGCAGCGAGGTCTATGACCAACATGTTGTTGAGAACATGCTGTGTGGCGGCATCCTC	1047
521	Db	ThrCysLysAlaSerThrLysValLysValThrAspAsnMetPheCysAlaGlyTyrSer	540
1048	QY	GGGAC-----CGCGAGGATCGCTCGAGGCGGCACAGTGGGGGGCCCATGGTCT---	1095
541	Db	ProGluAspSerLysArgGlyValAspLacysGluLysAspSerGlyLysProPheValMet	560

Qy	1096	--GCCCTCCATCACCAGCACTGGTTCCTGCGGTGGCCCTGGTGACTGGGGTAGGGCTGT	1155
Dd	561	LysSerProAspAspSerArgTrpTrpGlnValGlyIleValSerTrpGlyGluGlyCys	580
Qy	1153	GGGCTCCTTCACAACTACGGCGTTTACACCAAAGTCACGCCTACCTCGACTGGATC	1209
Dd	581	AspArgAspGlyLysTyrglyPheThrHisValPheArgLeuLysLysTrpIle	599
 RESULT 16 Q91001 PRELIMINARY; PRT; 607 AA.			
ID	Q91001		
AC	Q91001;		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	Thrombin.		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
OC	Gallus.		
OX	NCBI_TaxID=9031;		
[1]			
RN	SEQUENCE FROM N.A.		
RP	TISSUE=LIVER;		
RC	MEDLINE=92212913; PubMed=1557383;		
RX	Banfield D.K., MacGillivray R.T.;		
RA	"Partial characterization of vertebrate prothrombin cDNAs:		
RT	amplification and sequence analysis of the B chain of thrombin from		
FT	nine different species.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).		
[2]			
RN	SEQUENCE FROM N.A.		
RP	TISSUE=LIVER;		
RC	MEDLINE=94223694; PubMed=7513365;		
RX	Banfield D.K., Irwin D.M., Waiz D.A., MacGillivray R.T.;		
RA	"Evolution of prothrombin: isolation and characterization of the cDNAs		
RT	encoding chicken and hagfish prothrombin.";		
RL	J. Mol. Evol. 38:177-187(1994).		
[3]			
RN	SEQUENCE FROM N.A.		
RP	TISSUE=LIVER;		
RC	Banfield D.K.;		
RL	Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; M61391; AAA21619.1; -.		
DR	HSP; P00734; IUUS.		
DR	MEROPS; S01.217; -.		
DR	InterPro; IPR001314; Chymotrypsin.		
DR	InterPro; IPR002383; GLA_blood.		
DR	InterPro; IPR000001; Kringle.		
DR	InterPro; IPR003966; Prothrombin.		
DR	InterPro; IPR001254; Ser_protease_Try.		
DR	InterPro; IPR000294; VitK_dep_GLA.		
DR	Pfam; PF00594; gla; 1.		
DR	Pfam; PF00051; kringle; 2.		
DR	Pfam; PF00089; trypsin; 1.		
DR	PRINTS; PR00722; CHYMOTRYPSIN.		
DR	PRINTS; PR00001; GLABLOOD.		
DR	PRINTS; PR00018; KRINGLE.		
DR	PRINTS; PR01505; PROTHROMBIN.		
DR	ProDom; PD000395; Kringle; 2.		
DR	SMART; SM00069; GLA; 1.		
DR	SMART; SM00130; KR; 2.		
DR	SMART; SM00020; TRY_P_SPC; 1.		
DR	PROSITE; PS00011; GLU_CARBOXYLATION; 1.		
DR	PROSITE; PS00021; KRINGLE_1; 2.		
DR	PROSITE; PS00070; KRINGLE_2; 2.		
DR	PROSITE; PS02440; TRYPsin_DOM; 1.		
DR	PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.		
DR	PROSITE; PS00135; TRYPsin_SER; 1.		
SQ	Hydrolase; Serine protease.		
SW	SEQUENCE 607 AA; 69110 MW; 002F3606EA36270F CRC64;		



Alignment Scores:			
Pred. No.:	6.12e-29	Length:	799
Score:	510.50	Matches:	137
Percent Similarity:	48.0%	Conservative:	52
Best Local Similarity:	34.86%	Mismatches:	141
Query Match:	21.84%	Indels:	63
DB:	11	Gaps:	19
US-09-763-153-1 (1-1245) x Q9DBI0 (1-799)			
Qy	118	TTCTGTGTCGAAGCAGCTCGACGGTGACCAAGTGTGCTTGTCTTCCCTTGAGCACCCTGC	177
		:::::     :::	
Db	436	TyTyTySerLeuTyraSngInSerAspProCys-----	ProGly 448
Qy	178	GCCAGCCTTGCTGCGCGGCACGCACGTGCATC-----	GACGGCATCGCAGCTTC 228
		::::     ::	
Db	449	GlupheLeuCysSerValaSngLyLeuCysValProAlaCysAspGlyLeIlys----	466
Qy	229	AGCTGCGCACTGCCGACGCGCTGGGAGGCCGCTTCTTGCACGGCAGGTGAGCTTCCTC	288
		::::  ::::	
Db	467	-----AspCysProASngLyLeuAspGluAtrGAsnCysValCysArgAlaMetPhe---	483
Qy	289	AATTGCTCGCTGGACAACGGCGCTGCAGCACTTACTGCCTAGAGAGGTGGGTGGCGG	348
		:::	
Db	484	GlncCysGlnGluAspSer-----	ThrCysIleSerLeuProArgVal----- 497
Qy	349	CGCTGTAGCTGTGGCCTTGCTTACAAGCTGGGGGACGACCTCTCTCAGTGTCCACCCCGCA	408
		::	
Db	498	-----CysAspAraGlnProAspCysLeuASngLySerASpGluGluGlnCysGlnGluGly	516

RN SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC019376, AAH19376.1; -  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF\_2.  
 DR InterPro: IPR001438; EGF\_II.  
 DR InterPro: IPR000083; Fibronectn.  
 DR InterPro: IPR000582; FN\_Type\_II.  
 DR InterPro: IPR000001; Kringle.  
 DR InterPro: IPR001254; Ser\_protease\_Try.  
 DR Pfam: PF00008; EGF; 2.  
 DR Pfam: PF00039; fn1; 1.  
 DR Pfam: PF00040; fn2; 1.  
 DR Pfam: PF00051; kringle; 1.  
 DR Pfam: PF00089; trypsin; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PRINTS: PR00010; EGF\_BLOOD.  
 DR PRINTS: PR00013; FNTYPEII.  
 DR PRINTS: PR00018; KRINGLE.  
 DR ProDom: PD000395; Kringle; 1.  
 DR ProDom: PD000995; FN\_Type\_II; 1.  
 DR SMART: SM00181; EGF; 2.  
 DR SMART: SM00058; FN1; 1.  
 DR SMART: SM00059; FN2; 1.  
 DR SMART: SM00130; KR; 1.  
 DR SMART: SM00020; TRYP\_SPC; 1.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_2.  
 DR PROSITE: PS01186; EGF\_2; UNKNOWN\_1.  
 DR PROSITE: PS01253; FIBRONECTIN\_1; UNKNOWN\_1.  
 DR PROSITE: PS00023; FIBRONECTIN\_2; UNKNOWN\_1.  
 DR PROSITE: PS00021; KRINGLE\_1; UNKNOWN\_1.  
 DR PROSITE: PS00070; KRINGLE\_2; 1.  
 DR PROSITE: PS02040; TRYP\_SIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYP\_SIN\_HIS; UNKNOWN\_1.  
 DR PROSITE: PS00135; TRYP\_SIN\_SER; UNKNOWN\_1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Alignment Scores:  
 Pred. No.: 1,49e-26 Length: 653  
 Score: 477.50 Matches: 136  
 Percent Similarity: 41.60% Conservative: 67  
 Best Local Similarity: 27.87% Mismatches: 166  
 Query Match: 20.43% Indels: 119  
 DB: 11 Gaps: 19

US-09-763-153-1 (1-1245) x Q8VCS4 (1-653)

43 CGGGAGTCATAGAGGAGATCTGTGACTTCGAGGAGCCAAAGAAATTTTCCAAATGTG 102  
 192 LysAspCysGlyThrGluLysCys---PheAspGluThrArgTyrGluTyrPheGluVal 210  
 103 GATGACACACTGGCCTTCTGGTCCAGCAGCTGCGAGGT-----GACCATGTGTGTC 156  
 211 GlyAsp-----HisTrpAlaArgValSerGluGlyHisValGluGlnCysGlyCys 227  
 157 TTGCCCTTTGGAGCACCCCTGC-----GCCAGCCTGTGC 189  
 228 MetGluGlyGlnAlaArgCysGluAspThrHisHisThrAlaCysLeuSerSerProCys 247  
 190 TCGGGGACGGCACCTGC-----ATCGACGGCATCGGCAGCTTCAGCTGCGACTGCGCG 243  
 248 LeuAnGlyGlyThrCysHisLeuValGlyThrGlyThrSerValCysThrCysPro 267  
 244 ACAGCTGGAGGGCCGCTTCTGCGAGCGGAGGTGACCTTCATTAATTCCTCGCTGGAC 303  
 268 LeuGlyTyrAlaGlyArgPheCys---AsnIleValProThrGluHisCysPheLeuGly 286  
 304 AACGGC-----

Db 287 AsnGlyThrGluTyrArgGlyValAlaSerThrAlaAlaSerGlyLeuSerCysLeuAla 306  
 QY 309 ----- 309  
 Db 307 TrpAsnSerAspLeuLeuTyrGlnGluLeuHisValAspSerValAlaAlaValLeu 326  
 QY 310 -----GGCTGCACCATCTACTGCTA----- 330  
 Db 327 LeuGlyLeuGlyProHisAlaTyrCysArgAsnProAspLysAspGluArgProTrpCys 346  
 QY 331 -----GAGGAGTGGGTGGCGGCTGT-----AGC 357  
 Db 347 TyrValValLysAspAsnAlaLeuSerTrpGluTyrCysArgLeuThrAlaCysGluSer 366  
 QY 358 TGTGGCGCTGTGTACAGCTGGGAGAGCCTCTG-----CAGTGTACACCC 405  
 Db 367 LeuAlaArgValHisSerGlnSerProGluIleLeuAlaAlaLeuProGluSerAlaPro 386  
 QY 406 GCAGTGAAGTTCCCTGTGGGAGGCCCTGGAGCGGATGGAGAGAGCAGCAGTCACCTG 465  
 Db 387 AlaValArgProThrCysGly-----LysArgHisLysLysArgThrPheLeu 402  
 QY 466 AAACGAGACACAGAGACCAAGAACCAAGACCTAGTACCGGGCTCATTTGATGGGAAGATG 525  
 Db 403 Arg-----ProargIleIleGlyGlySerSer 411  
 QY 526 ACCAGGGGGGAGACAGCCCTGTCAGTGTGTCTGTGGACTCAAGAGAGAGCTGCC 585  
 Db 412 SerLeuProGlySerHisProTrpLeuAlaAlaIleTyrIleGlyAsnSer-----Phe 429  
 QY 586 TCGGGGCGAGTCTCATCCACCCCTCTGCTGGTGTGTGACAGCGGCCACTCATGATGATGAG 645  
 Db 430 CysAlaGlySerLeuValHisThrCysTrpValSerAlaAlaHisCysPheAlaAsn 449  
 QY 646 TCC-----AAGAAGCTCTTGTGAGGTGGAGAGTATGACCTGCGCGCTGGGAG 696  
 Db 450 SerProArgAspSerIleThrValValLeuGlyGlnHisPheAsnAlaThrThr 469  
 QY 697 AAGTGGGAGCTGGACCTGACCAAGAGGTCTTCTCCACCCCACTCAAGAGAGCAGG 756  
 Db 470 AspValThrGlnThrPheGlyIleGluTyrValProTyrThrLeuTyrSerValPhe 489  
 QY 757 ACCACCGACAAT---GACATCGCACTGCTGCACCTGCCAGCCAGCC-----GCC 801  
 Db 490 AsnProAsnAsnHisAspLeuValIleArgLeuLysLysLysGlyGluArgCysAla 509  
 QY 802 ACCCTCTCGCAGACCATAGTCCCATCTGCTCCCGAGACAGCGCTTCGACAGCGCGAG 861  
 Db 510 ValArgSerGlnPheValGlnProIleCysLeuProGluAlaGlySerSer----- 526  
 QY 862 CTCAATCAGCGCCGAGGAGACCCCTGTCAGCGGTGGGGTACCACAGCAGCAGCGAGAG 921  
 Db 527 ---PheProThrGlyHisLysCysGlnIleAlaGlyTrpGlyHisMetAspGluAsnVal 545  
 QY 922 AAGGAGCCCAAGAGAAACCCACCTTGTCTCTCAACTTCATCAAGATTCCCGTGGTCCCG 981  
 Db 546 SerSerTyrSerAsnSerLeuLeuGluAlaLeu-----ValProLeuValAla 561  
 QY 982 CACAATCAGTGCAGC-----GAGGTCAATCAGCAACATGCTGTCTGAGACATGCTGTGT 1035  
 Db 562 AspHisLysCysSerProGluValTyrGlyAlaAspIleSerProAsnMetLeuCys 581  
 QY 1036 CGCGGATCTCTCGGGAGGAGGATGCTGCGAGGGGACAGTGGGGGGCCCATGGTC 1095  
 Db 582 AlaGlyTyrPheAspCysLysSerAspAlaCysGlnGlyAspSerGlyGlyProLeuVal 601  
 QY 1096 GCCTCTTCACGGCAGCTGTTCTTCTGGGGCTTGGTGGGTGGGTGGGTGGGTGGGTGG 1155  
 Db 602 CysGluLysAsnGlyValAlaTyrLeuTyrGlyIleIleSerTrpGlyAspGlyCysGly 621  
 QY 1156 CTCTCTCACAACTACCGCGTTTACCAAGTACAGCGGTACCTCGACTGGATCCATGGG 1215

Db 622 ArgLeuAsnLysProGlyValThrArgValAlaAsnTyrValAspTyrIleAsnAsp 641  
 QY 1216 CACATCAGACAAAGGAGCCGCC 1239  
 Db 642 ArgIleArgProLysArgPro 649  
 RESULT 19  
 Q9CQW3  
 ID Q9CQW3 PRELIMINARY; PRT; 399 AA.  
 AC Q9CQW3  
 DT 01-JUN-2001 (TRENBLrel. 17, Created)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE 1300015B06Rik protein.  
 1300015B06Rik.  
 Mus musculus (Mouse).  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=STOMACH, AND LIVER;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 Hayashizaki Y.;  
 "Functional annotation of a full-length mouse cDNA collection."  
 Nature 409:685-690(2001).  
 RL Nature 409:685-690(2001).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 TRYPSIN FAMILY.  
 EMBL; AK008819; BAB25912.1; -.  
 EMBL; AK005011; BAB23744.1; -.  
 HSP; P00740; ICFH.  
 MEROPS; S01.996; -.  
 MGD; MGI:1914151; 1300015B06Rik.  
 InterPro; IPR000152; Asx\_hydroxyl.  
 InterPro; IPR001314; Chymotrypsin.  
 InterPro; IPR000561; EGF-like.  
 InterPro; IPR000742; EGF\_2.  
 InterPro; IPR001881; EGF\_Ca.  
 InterPro; IPR002383; GLA\_blood.  
 InterPro; IPR001254; Ser\_protease\_Try.  
 InterPro; IPR000294; vitk\_dep\_GLA.  
 Pfam; PF00008; EGF; 2.  
 Pfam; PF00594; gla; 1.  
 Pfam; PF00089; trypsin; 1.  
 PRINTS; PR00722; CHYMOTRYPSIN.  
 PRINTS; PR00001; GLABLOOD.  
 SMART; SM00181; EGF; 2.  
 SMART; SM00179; EGF\_CA; 2.  
 SMART; SM00069; GLA; 1.  
 SMART; SM00020; Tryp\_SPC; 1.  
 PROSITE; PS00010; ASX\_HYDROXYL; UNKNOWN\_1.  
 PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 PROSITE; PS00186; EGF\_2; 2.  
 PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
 PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 Hydrolase; Serine protease.

SQ SEQUENCE 399 AA; 44304 MW; 4FC85C9598F27E03 CRC64;  
 Alignment Scores:  
 Pred. No.: 4,85e-26 Length: 399  
 Score: 470.00 Matches: 126  
 Percent Similarity: 43.71% Conservative: 58  
 Best Local Similarity: 29.93% Mismatches: 143  
 Query Match: 20.11% Indels: 94  
 DB: 11 Gaps: 16  
 US-09-763-153-1 (1-1245) x Q9CQW3 (1-399)  
 QY 10 TTCCTGGAGGAGCTCCGTACAGCAGCGCTGGAGCGGAGTGCATAGAGAGATCTGTGAC 69  
 Db 45 PheLeuGluGluIlePheGlnGlyAsnLeuGluLysGluCysTyrGluGluValCysAsn 64  
 QY 70 TTCGAGGAGGCAAGGAAATTTCCAAATGTGATGACACACTGGCTTCTGTGTCAG 129  
 Db 65 TyrGluGluAlaArgGluValPheGluAsnAspValIleThrAspGluPheTyrArgGln 84  
 QY 130 CAGCTCGAGCGTGCACCATGCTTGTGCTTGCCTTGGAGCACCCTGCGCAGCCTGTGC 189  
 Db 85 TyrGlyGlyGlySerProCysValSerGlnPro-----Cys 96  
 QY 190 TCGGGGACGACGCTGCATCGCAGCGCTGGAGTTCAGCTGCGACTGCGCAGCGCGC 249  
 Db 97 LeuAsnGlnGlyThrCysGluAspHisIleArgSerTyrSerCysThrCysSerProGly 116  
 QY 250 TGGAGGCGCGCTTCTGCCAGCGGAGTGTGCTTCTCAATGCTCGCTGAC---AAC 306  
 Db 117 TyrGluGlyLysThrCysAlaMetAlaLysAsn-----GluCysHisLeuGluArgThr 134  
 QY 307 GCGCGCTGCAGCATCTACTGCTAGAGAGGTGGCTGGCGGCTGTAGCTGTGCGCCT 366  
 Db 135 AspGlyCysGlnHisPheCysHisProGlyGlnSerSerTyrMetCysSerCysAlaLys 154  
 QY 367 GGCTFACAGCTGGGGGAGGACCTCTTCAGTGTACACCGCGAGTGAAGTTCCTGTGGG 426  
 Db 155 GlyTyrLysLeuGlyLysAspGlnLysSerCysGlyProSerAspLysCysAlaCysGly 174  
 QY 427 AGCCCTTGAAGCGGATGGAGAGAGCGAGTCACTGAAACGAGACACACAGAACCA 486  
 Db 175 -----AlaLeuThrSerGluHisIle----- 181  
 QY 487 GAAGACCAAGTAGATCGCGGCTCATTTGATGGAAGATACACAGCGGGGAGACAGC--- 543  
 Db 182 -----ArgMetThrLysSerSerGlnSerGln 190  
 QY 544 -----CCCTGGCAGGTGCTCTGCTGGACTCAAGAGAAAGAGCTGCGCTGCGGGCA 594  
 Db 191 ProSerPheProTyrGlnValArgLeuThrAsnSerGluGlyGluAspPheCysAlaGly 210  
 QY 595 GTCTCATCCACCCCTCTCGGTGTGACAGCGCGCCACTGCATGGATGAGTCCCAAGAAG 654  
 Db 211 ValLeuGlnGluAspPheValLeuThrThrAlaLysCysSerLeuLeuHisSerAsn 230  
 QY 655 CTCTTGTTCAGCTTGGAGAGTATGACCTGCGCGCTGGAGAGTGGAGAGTGGAGCTG 714  
 Db 231 IleSerValLys---AlaAsnValAspGlnArg-----Ile 241  
 QY 715 GACATCAAGAGGTCTTCTGTCACCCCACTACACAAAGAGACACCCAGCAATGACATC 774  
 Db 242 ArgIleLysSerThrHisValHisMetArgTyrAspGluGluSerGlyGluAsnAspVal 261  
 QY 775 GCATGCTGCACCTGGCCAGCGCCGCCACCTCTCTGCGAGACCATAGTGCCTGCTC 834  
 Db 262 SerLeuLeuGlnLeuGluGluProLeuGlnCysProSerSerGlyLeuProValCysVal 281  
 QY 835 CCGGACAGCGGCTTTCAGAGCGCGGAGCTCAATCAGCGCGCCAGGAGACCTCGTCAAG 894  
 Db 282 ProGluArgAspPheAlaGluHisValLeu---IleProGlyThrGluGlyLeuLeuSer 300  
 QY 895 GGCTGGGCTACACAGCAGCGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 954

```

Db 301 GlyTrp-----MetLeu 304
QY 955 AACTTCATCAAGATCCCGTGGTCCG-----CACAA 987
Db 305 AsnGlyThrHisLeuAlaThrThrProMetLeuLeuSerValThrGlnAlaAspGlyGlu 324
QY 988 GAGTGCAGCAGGAGTGCATGAGCAACATGGTCTGTGAGAACATCTGTGT-----GCG 1038
Db 325 GluCysGlyGlnThrLeuAsnValThrValThrArgThrSerCysGluGlySer 344
QY 1039 GGCATCTCGGACCGGAGGAGTGCCTCGGAGGGGACAGTGGGGCCCATGTCGCGC 1098
Db 345 ValValMetGlyProTrpValGlu-----GlySerValValThr 357
QY 1099 TCCATTCACAC---GGCACCTGTGCTTCTGGTGGGCTGGTGGAGCTGG-----GGT 1143
Db 358 ArgGluHisIlysGlyThrTrpPheLeuThrGlyLeuLeuGlySerProProProGly 377
QY 1144 GAGGCTGTGGGCTCTTCACACTAGCGGCGTTTACACCAAGTACGCGCTACCTCGAC 1203
Db 378 GlnSerGlnMetLeuLeu-----LeuThrAlaValProArgTyrSerMet 392
QY 1204 TGG 1206
Db 393 Trp 393

RESULT 20
Q9V4N6 PRELIMINARY; PRT; 1379 AA.
AC Q9V4N6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE C2G2105 protein.
GN Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Aghavani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houlston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Moberg C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

```

```

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers F.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC EMBL: AE003840; AAF59230.1; -.
CC HSP: P00763; IDPO.
CC FlyBase: Fgn0033192; CG2105.
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR000024; Fz_domain.
CC InterPro: IPR002172; LDL_recept_A.
CC InterPro: IPR001254; Ser_protease_Try.
CC InterPro: IPR001190; Srcr_receptor.
CC InterPro: IPR000217; Tubulin.
CC Pfam: PF01392; Fz; 1.
CC Pfam: PF00057; trypsin; 1.
CC Pfam: PF00089; trypsin; 1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC PRINTS: PR00261; LDLRECEPTOR.
CC SMART: SM00063; FRI; 1.
CC SMART: SM00192; LDLa; 2.
CC SMART: SM00202; SR; 1.
CC SMART: SM00020; Tryp_SPC; 1.
CC PROSITE: PS00038; Fz; 1.
CC PROSITE: PS01209; LDLRA_1; 1.
CC PROSITE: PS00068; LDLRA_2; 2.
CC PROSITE: PS00287; SRCR_2; 1.
CC PROSITE: PS00240; TRYPSIN_DOM; 1.
CC PROSITE: PS00134; TRYPSIN_HIS; 1.
CC PROSITE: PS00135; TRYPSIN_SER; 1.
CC PROSITE: PS00227; TUBULIN; 1.
KW Glycoprotein; Hydrolase; Serine protease.
SQ SEQUENCE 1379 AA; 149489 MW; A393A9CC2167E4B7 CRC64;

Alignment Scores:
Pred. No.: 8,34e-24 Length: 1379
Score: 440.50 Matches: 133
Percent Similarity: 39.76% Conservative: 63
Best Local Similarity: 26.98% Mismatches: 174
Query Match: 18.85% Indels: 123
DB: 5 Gaps: 20

US-09-763-153-1 (1-1245) x Q9V4N6 (1-1379)
QY 55 GAGGAGATCTGTGACTTCGAGGAGCCCAAGAAATTTCCAAAATGTGGATGACACACTG 114
Db 872 GluAspCysValGlyLeuAspGluValArgGluValMetArg----- 885
QY 115 GCCTTCTGTGTCACAGCAGCTGCAGCGTGCACGATGC-----TTGGTTCCTCCCTTG 165
Db 886 AlaAlaThrHisProLysCysAspGlyPheGlnCysAspGlnAsnArgCysLeuProGln 905
QY 166 GAGCACCCCTGTC----- 177
Db 906 GluTyrValCysAspGlyHisLeuAspCysMetAspGlnAlaAspGluAlaLysCysGlu 925
QY 178 -----GCCAGCTGTGTCGCGGACGCGCAGCTGCATC----- 210
Db 926 ArgCysGlyProAspGluIleThrCysGlyAspSerGlnCysIleGlyThrLysHisIle 945
QY 211 ---GAGCGAGTCGCGAGCTTCAGCTCCGACTCCGCGCAGCGCTGGAGGGCGCTTCTGCG 267
Db 946 CysAspGlyIle-----IleAspCysProTyrGlyGlnAspGluArgAsnCys 961
QY 268 CAGCGGAGTGTGAGTTCCTCAATTCTCGTGGACACGCGGCTGCACGCACTTACTGCG 327

```

Db 962 LeuArg---LeuSerGluArgAsnGlyAspValGlyThrGlyValLeuGluValTyrArg 980  
QY 328 CTAGAGAGGTGGCTGGCGCGCTAGCTGCTGGCCCTGCTACAAG----- 375  
Db 981 IleGlyGlnArgGlnTrpMetPro---AlaCysValLysAsnTrpAspArgAlaValSer 999  
QY 376 -----CTGGGGACGACCTCTGCAGTGTCAACCCCGCAGTGAAG 414  
Db 1000 ProSerAlaValCysSerIleLeuGlyTyrSerAlaValAsnAlaThrSerValLeuThr 1019  
QY 415 TTCCCTTGTGGAGGCC-----TGAAGCGG 441  
Db 1020 GlnLeuThrHisArgProGluLeuAlaThrValAsnValSerThrAspIleTrpLysMet 1039  
442 ATGGAGAAAGAGCGAGTCACTCACTGAACGACAGAC-----ACAGAAGAC 483  
1040 TyrAlaLysArgLysSerThrLeuMetGlnGluPheAlaAsnCysLysLysThrGluAsp 1059  
QY 484 CAA-----GAA 489  
Db 1060 TyrProMetAlaAspLeuThrCysSerAsnTyrGluCysGlyArgValLysArgGlyArg 1079  
QY 490 GACCAAGTAGATCCGCGCTCATTTGAGGGAAGATGACCGCGGGAGACAGCCCTGG 549  
Db 1080 HisLysProSerArgArgIleIleGlyThrGlnAlaSerProGlyAsnTrpProPhe 1099  
QY 550 CAGGTGGTCTGCTGAGTCAAGAAAGAGCTG---GCCTCGGGGGCAGTGTCTCATCCAC 606  
Db 1100 LeuAlaIleLeuGlyGlyProGluLysIlePheTyrCysAlaGlyValLeuIleSer 1119  
QY 607 CCTCTCTGGTGTGACAGCGGCCCTGCTGATGATGAGTCCCAAGAGCTCTGTTCAGG 666  
Db 1120 AspGlnTrpValLeuThrAlaSerHisCys----- 1129  
QY 667 CTGGAGAGTATGACCTGCGCGCTGGGAGAGTGGAGCTGGACCTG----- 714  
Db 1130 ValGlyAsnTyrSerValIleAspLeuGluAspTrpThrIleGlnLeuGlyValThrArg 1149  
QY 715 -----GACATCAAGGAGGTCTTCGTCCACCC 741  
Db 1150 ArgAsnSerPheThrTyrSerGlyGlnLysValLysValLysAlaValIleProHisPro 1169  
QY 742 AACTACAGACAGACACACCC---GACAACTACATCGCTGCTGACCTGGCGCCAGCC 798  
Db 1170 GlnTyrAsnMetAlaIleAlaHisAspAsnAspIleAlaLeuPheGlnLeuAlaThrArg 1189  
QY 799 GCCACCCTCTCGCAGACCATAGTCCCTGCTGCTCCGACAGCGCCCTTGCAGAGCGC 858  
Db 1190 ValAlaPheHisGluHisLeuLeuProValCysLeuProProSerVal-----Arg 1207  
QY 859 GAGCTCAATCAGCGCGCCAGGAGACCCCTGCTGACGGCTGGGCTACACAGCAGCGCA 918  
Db 1208 AsnLeuHis---ProGlyThrLeuCysThrValIleGlyTrpGly-----LysArgGlu 1224  
QY 919 GAGAGGAGGCGCAGAGAAACCGACCTTCTGCTCATCAATTCATCAAGATTCCTGGTC 978  
Db 1225 AspLysAspProLysSerThrTyrGluTyrIleValAsnGluValGlnValProIleIle 1244  
QY 979 CCGCACAAATGAGTGCAGCGAGTCAAGCAACATG---GTGCTCAGAACATGCTGTGT 1035  
Db 1245 ThrArgAsnGlnCysAspGluTrpLeuAspAsnLeuThrValSerGluGlyMetValCys 1264  
QY 1036 GCGGGCATCTCGGGACCGCAGGATGCTGCGAGGGGACAGTGGGGGGCCCATGGTC 1095  
Db 1265 AlaglyPheAspAspGlyGlyLysAspAlaCysGlnGlyAspSerGlyGlyProLeuLeu 1284  
QY 1096 GCCTCTTCCACGGC-----ACCTGTTCTGGTGGCTGTGTGAGTGGGTGAG 1146  
Db 1285 CysProTyrProGlyGluLysAsnArgTrpPheValGlyIleValSerTrpGlyIle 1304  
QY 1147 GGCTGTGGTCTTCACTACAGGCTTACACCAAGAGTCCAGCGCTACCTCGACTGG 1206  
Db 1305 MetCysAlaHisProArgLeuProGlyValTyrAlaAsnValGlnTyrValProTrp 1324

QY 1207 ATCCATGGGCACATFCAGAGACAAGGAGCCCGCCAGAG 1245  
Db 1325 IleglnGluGlnIleAlaLysHisSerArgProIleLys 1337

Search completed: March 17, 2003, 17:35:20  
Job time : 95.5 secs





GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: March 17, 2003, 17:28:54 ; Search time 38 Seconds  
(without alignments)  
1927.976 Million cell updates/sec

Title: US-09-763-153-1

Perfect score: 2337

Sequence: 1 gccaaactctctctggagga.....acaaggagagccccccaggaag 1245

Indexing table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 60 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool/US09763153/runat\_11032003\_084248\_2035/app\_query.fasta\_1.1415  
-DB=IssuedPatents\_AA -QPMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi  
-LIST=60 -DOALIGN=200 -THR\_SCORE=ptc -THR\_MAX=100 -THR\_MIN=0 -ALIGN=20  
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09763153 -ACGN\_1\_1\_21 @runat\_11032003\_084248\_2035 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pcp.\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pcp.\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pcp.\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pcp.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pcp.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2298	98.3	419	1	US-08-295-411-1
2	2298	98.3	419	2	US-08-955-471-1
3	2298	98.3	419	5	US-09-667-570A-3
4	2298	98.3	419	5	PCT-US92-10242-1
5	2298	98.3	460	2	US-08-756-506-2
6	2298	98.3	460	2	US-08-756-506-4
7	2298	98.3	461	6	US-08-525537-2
8	2292	98.1	461	6	5460953-3
9	2266	97.0	461	6	5270178-17
10	2266	97.0	461	6	5270178-18
11	2260.5	96.7	460	6	5270178-13
12	2260.5	96.7	460	6	5270178-14

13	2255	96.5	410	4	US-09-065-872-1	Sequence 1, Appli
14	2255	96.5	410	4	US-09-667-570A-1	Sequence 1, Appli
15	2253.5	96.4	461	6	5270178-2	Patent No. 5270178
16	2247	96.1	409	4	US-09-065-872-2	Sequence 2, Appli
17	2247	96.1	409	4	US-09-667-570A-2	Sequence 2, Appli
18	2245.5	96.1	460	6	5270178-15	Patent No. 5270178
19	2218.5	94.9	460	6	5270178-16	Patent No. 5270178
20	1393	59.6	262	1	US-07-720-189-1	Sequence 1, Appli
21	1367.5	58.5	261	6	5270178-19	Patent No. 5270178
22	1358.5	58.1	261	6	5270178-20	Patent No. 5270178
23	1328	56.8	250	4	US-08-944-483-51	Sequence 51, Appli
24	1320.5	56.5	261	6	5270178-21	Patent No. 5270178
25	1302.5	55.7	261	6	5270178-5	Patent No. 5270178
26	809	34.6	487	1	US-08-469-486-53	Sequence 53, Appli
27	809	34.6	487	2	US-08-469-486-53	Sequence 53, Appli
28	809	34.6	492	1	US-08-469-486-2	Sequence 2, Appli
29	809	34.6	492	2	US-08-469-486-2	Sequence 2, Appli
30	801.5	34.3	448	5	PCT-US92-10068-1	Sequence 1, Appli
31	799.5	34.2	448	1	US-08-295-411-3	Sequence 3, Appli
32	799.5	34.2	448	1	US-08-295-411-3	Sequence 3, Appli
33	799.5	34.2	448	5	PCT-US92-10242-3	Sequence 3, Appli
34	783	33.5	406	1	US-08-955-471-5	Sequence 5, Appli
35	783	33.5	406	2	US-08-955-471-5	Sequence 5, Appli
36	783	33.5	406	5	PCT-US92-10242-5	Sequence 5, Appli
37	783	33.5	444	1	US-08-475-845-2	Sequence 2, Appli
38	783	33.5	444	2	US-08-327-690-2	Sequence 2, Appli
39	783	33.5	444	2	US-08-660-289-2	Sequence 2, Appli
40	783	33.5	444	2	US-08-537-807-2	Sequence 2, Appli
41	783	33.5	444	2	US-08-871-003-2	Sequence 2, Appli
42	783	33.5	444	3	US-08-464-233-2	Sequence 2, Appli
43	783	33.5	444	4	US-09-189-607-2	Sequence 2, Appli
44	783	33.5	444	4	US-09-378-907-2	Sequence 2, Appli
45	783	33.5	444	5	PCT-US94-05779-2	Sequence 2, Appli
46	783	33.5	466	1	US-07-882-202A-4	Sequence 4, Appli
47	783	33.5	466	1	US-08-021-615A-4	Sequence 4, Appli
48	783	33.5	466	1	US-08-321-777-4	Sequence 4, Appli
49	783	33.5	466	4	US-09-009-217-14	Sequence 14, Appli
50	783	33.5	466	4	US-09-009-656-14	Sequence 14, Appli
51	783	33.5	466	5	PCT-US93-04493-4	Sequence 4, Appli
52	742	31.8	415	1	US-08-073-531B-1	Sequence 1, Appli
53	742	31.8	415	2	US-08-766-288-1	Sequence 2, Appli
54	742	31.8	437	1	US-08-487-037-2	Sequence 2, Appli
55	738.5	31.6	488	1	US-08-487-037-1	Sequence 1, Appli
56	736.5	31.5	461	6	5521070-2	Patent No. 5521070
57	736	31.5	461	3	US-08-742-877-2	Sequence 2, Appli
58	736	31.5	461	4	US-09-053-871A-21	Sequence 21, Appli
59	734	31.4	437	1	US-08-487-037-3	Sequence 3, Appli
60	731	31.3	415	1	US-08-295-411-2	Sequence 2, Appli

#### ALIGNMENTS

RESULT 1  
US-08-295-411-1  
; Sequence 1, Application US/08295411  
; Patent No. 5679639

GENERAL INFORMATION:

APPLICANT: Griffin, John H.

INVENTOR: Westers, Rolf M.

TITLE OF INVENTION: Serine Protease-Derived Polypeptides and

ADDRESSEE: Anti-Peptide Antibodies, Systems and Therapeutic Methods

TITLE OF INVENTION: for Inhibiting Coagulation

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESS: Office of Patent Counsel, The Scripps

ADDRESS: Research Institute

STREET: 10666 No. 5679639th Torrey Pines Road, TPC 8

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/295,411  
FILING DATE: 22-AUG-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/793,989  
FILING DATE: 18-NOV-1991  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI263.0C1  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 419 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..157  
OTHER INFORMATION: /note= "Protein C Light Chain"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 158..169  
OTHER INFORMATION: /note= "Protein C Activation"  
OTHER INFORMATION: Peptide"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 170..419  
OTHER INFORMATION: /note= "Protein C Heavy Chain"  
US-08-295-411-1

Alignment Scores:  
Pred. No.: 3,34e-167 Length: 419  
Score: 2298.00 Matches: 415  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.33% Indels: 0  
DB: 1 Gaps: 0

US-09-763-153-1 (1-1245) x US-08-295-411-1 (1-419)

QY 1 GCCAACTCCTCTGGAGGAGCTCGTCCACAGCAGCCTGGAGCGGAGTGCATAGAGGAG 60  
|||||  
1 AlaasnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 20  
61 ATCTGTGACTTCGAGAGGCCAAGAAATTTCCAAAATGTGGATGACACACTGGCCCTC 120  
21 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspThrLeuAlaPhe 40  
QY 121 TGGTCCAGCAGCTGCAGCGGTGACAGAGTCTTGGTCTTCCCTTGGAGCACCCTGGCC 180  
|||||  
41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60  
QY 181 AGCGTGTGCTGGCGGAGCGGACGTGCATCGACGCGCATCGGCGAGCTTCAGCTGGCAGTGC 240  
|||||  
61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
QY 241 CGCAGCGGTGGAGCGCGCTTTCGCCAGCGCGAGGTGAGCTTCCTCAATGTCTCGCTG 300  
|||||  
81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
QY 301 GACACGGCGGCTGCAGCAGTACTACCTAGAGAGGTGGCTGCGCGCTGTAGCTGT 360  
|||||

Db 101 AspAsnGlyGlyCysThrHisTyrCysLeuGluGluValGlyTrpArgArgCysSerCys 120  
QY 361 GCGCCTGGCTACAAGCTGGGGAGGACACCTCTGCAGTGTACCCCGCAGTGAAGTCCCT 420  
|||||  
Db 121 AlaProGlyTyrLysLeuGlyAspSerLeuGlnCysHisProAlaValLysPhePro 140  
QY 421 TGTGGAGGCCCTGGAGCGGATGGAGAAGCGCAGTCACTGAAACGAGACACAGAA 480  
|||||  
Db 141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160  
QY 481 GACCAAGAAGACCAAGTATAGTCCGGGCTCATGTATGGGAAGATGACACAGCGGGAGAC 540  
|||||  
Db 161 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 180  
QY 541 AGCCCTGGCAGGTGCTCTGCTGGGACTCAAAGAAGAGCTGGCCCTCGGGGAGTGTCTC 600  
Db 181 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 200  
QY 601 ATCCACCCCTCTGCTGGTGTGCTGACAGCGGCCACTGCATGGATGAGTCCAAGAGCTCTT 660  
Db 201 IleHisProSerTrpValLeuThrAlaHisCysMetAspGluSerLysLeuLeu 220  
QY 661 GTCAGGCTTGGAGACTATGACCTGGCGGCTGGGAGAAGCTGGGAGCTGGACCTGGACATC 720  
Db 221 ValArgLeuGlyGluTyrAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 240  
QY 721 AAGGAGGTCTTCCTCCACCCCAACTACAGCAAGAGCACCACCAATGACATCGCAGCTG 780  
Db 241 LysGluValPheValHisProAsnTyrSerLysSerThrAspAsnAspIleAlaLeu 260  
QY 781 CTCACCTGGCCCGCCAGCCCGCCCTCTCGAGACCATAGTGCCTCATCTGCCCTCCCGGAC 840  
Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 280  
QY 841 AGCGGCTTTCAGAGCGGAGCTCAATCAGCGCGGCGGAGAGACCTCTGTCAGCGGCTGG 900  
Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
QY 901 GGCTACACAGCAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
Db 301 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
QY 961 ATCAAGATTCCTGGTGGTCCCGGCAATGAGTGCAGCGAGGTGCATGACCAATGTGTCT 1020  
Db 321 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
QY 1021 GAGAACATGCTGTGCGGCGCATCTCGGGGACCGGCGGAGGATCGCTCGGAGGCGGACGT 1080  
Db 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
QY 1081 GGGGGGCGCATGTGCGCTCTTCACCGGACCTGTGTCTTGTGGGCTGTGTGAGCTGG 1140  
Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
QY 1141 GGTGAGGCTGTGGGCTCTTCACACTAGCGGCTTACACCAAGTACAGCCGCTACCTC 1200  
Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 400  
QY 1201 GACTGGATCCATGGGACATCAGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245  
Db 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415  
RESULT 2  
US-08-955-471-1  
Sequence 1, Application US/08955471  
Patent No. 5968751  
GENERAL INFORMATION:  
APPLICANT: Griffia, John H.  
APPLICANT: Mesters, Rolf M.  
TITLE OF INVENTION: Serine protease-Derived Polypeptides and  
TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods  
TITLE OF INVENTION: for Inhibiting Coagulation  
NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Office of Patent Counsel, The Scripps  
 ADDRESSEE: Research Institute  
 STREET: 10666 No. 5968751th Torrey Pines Road, TPC 8  
 CITY: La Jolla  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 92037

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/955,471  
 FILING DATE:

CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/295,411  
 FILING DATE:

CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fitting, Thomas  
 REGISTRATION NUMBER: 34,163  
 REFERENCE/DOCKET NUMBER: TSRI263.0C1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 619-554-2937  
 TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 419 amino acids  
 TYPE: amino acid

TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO

NAME/KEY: Region  
 LOCATION: 1..157  
 OTHER INFORMATION: /note= "Protein C Light Chain"

NAME/KEY: Region  
 LOCATION: 158..169  
 OTHER INFORMATION: /note= "Protein C Activation  
 OTHER INFORMATION: Peptide"

NAME/KEY: Region  
 LOCATION: 170..419  
 OTHER INFORMATION: /note= "Protein C Heavy Chain"

US-09-763-153-1

# Alignment Scores:

Pred. No.:	3,34e-167	Length:	419
Score:	2298.00	Matches:	415
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.33%	Indels:	0
DB:	2	Gaps:	0

US-09-763-153-1 (1-1245) x US-08-955-471-1 (1-419)

QY	1	GCACACTCTCTCGAGAGCTCCGTCACAGCAGCTGGAGCGGAGTCATAGAGCAG	60
Db	1	AlaAsnSerPheLeuGluGluLeuArgHisSerLeuGluArgGluCysIleGluGlu	20
QY	61	ATCTGTGACTTCGAGAGGCCAAGAAATTTCCAAAATGTGGATGACACACTGGCCCTC	120
Db	21	IleCysAspPheGluAlaLysGluIlePheGlnAsnValAspThrLeuAlaPhe	40
QY	121	TGTCACACAGCTGCAGGTCACACAGTGCCTGGTCTTCCCTTGGAGCACCCTGGCC	180
Db	41	TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCys	60

QY	181	AGCCTGTCTCGCGGCACGGCAGCTGTCATCGACGCGCATCGCAGCTTCAGTGCAGCTGC	240
Db	61	SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys	80
QY	241	CGCAGCGCTGGAGAGCGCGCTTCTGCCAGCGCGAGGTGAGCTTCCTCAATGTCTGCTG	300
Db	81	ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu	100
QY	301	GACAACGGCGCTGCACGCATTACTGCTAGAGAGGTGGCTGGCGCGCTGTAGCTGT	360
Db	101	AspAsnGlyGlyCysThrHisThrCysLeuGluGluValGlyTrpArgCysSerCys	120
QY	361	CGCGCTGCTACAAGCTGGGGACGACCTCTCGCAGTGTCCACCCCGCAGTCAAGTTCCT	420
Db	121	AlaProGlyTyrLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro	140
QY	421	TGTGGAGCGCTGGAAGCGGATGGAGAAGCGCAGTCACCTGAAACGAGACACAGAA	480
Db	141	CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu	160
QY	481	GACCAAGAAGACCAAGTAGATCCGCGCTCATTCATGGAGATGACACGCGGGGAGAC	540
Db	161	AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp	180
QY	541	AGCCCTCGCGAGGTGCTCTGCTGACTCAAGAAGAGCTGGCTCGCGGCGAGTCTC	600
Db	181	SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu	200
QY	601	ATCCACCCCTCTCGGTGTGACAGCGGCCACTGTCATGGATGAGTCCAGAGAGCTCCT	660
Db	201	IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu	220
QY	661	GTCAGGCTTGAGAGTATGACCTCGCGCGCTGGAGAGTGGAGCTGGACCTGGACATC	720
Db	221	ValArgLeuGlyGluTyrAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle	240
QY	721	AAGGAGTCTTCGTCCACCCCACTACAGCAAGAGCACCACCGCAATGATCGCAGCTG	780
Db	241	LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu	260
QY	781	CTGCACTGCGCCAGCGCCGACCTCTCGCAGACCACATAGTGCCTTCCTCCCGGAC	840
Db	261	LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProLysLeuProAsp	280
QY	841	AGCGCGCTTCGAGAGCGGAGCTCAATCAGCGCGCGCAGGAGCCCTCGTGCAGGGCTGG	900
Db	281	SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp	300
QY	901	GGCTACCAACAGCAGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	960
Db	301	GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe	320
QY	961	ATCAAGATTCGCGGTCCCGCACAAATGATGTCAGCAGAGTGCATGACCAACATGGTCTCT	1020
Db	321	IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer	340
QY	1021	GAGAACATGTGTGTGCGGGCATCTCGGGGACCGGAGGAGTGCCTCGGAGGGGCGACGT	1080
Db	341	GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer	360
QY	1081	GGGGGGCCATGTCGCTCTCTCCAGCGCACCTGCTTCTGTTGGTGGCTGGTGGAGTGG	1140
Db	361	GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp	380
QY	1141	GGTGAGGCTGTGGGCTCTTCACAACTACGGGCTTTTACACCAAGTACGCGCTACCTC	1200
Db	381	GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu	400
QY	1201	GACTGGATTCATGGGCACATCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1245
Db	401	AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys	415

RESULT 3

## US-09-667-570A-3

; Sequence 3, Application US/09667570A

; Patent No. 6436397

## ; GENERAL INFORMATION:

; APPLICANT: Baker, Jeffrey C

; APPLICANT: Carlson, Andrew D

; APPLICANT: Huang, Lihua

; APPLICANT: Sheliga, Theodore A

; TITLE OF INVENTION: Improved Methods for Processing Activated Protein C

; FILE REFERENCE: X-11796A

; CURRENT APPLICATION NUMBER: US/09/667.570A

; CURRENT FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: 60/045,255

; PRIOR FILING DATE: 1997-04-28

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 419

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-667-570A-3

## ; Alignment Scores:

; Seq. No.: 3, 34e-167 Length: 419

; Score: 2298.00 Matches: 415

; Percent Similarity: 100.00% Conservative: 0

; Best Local Similarity: 100.00% Mismatches: 0

; Query Match: 98.33% Indels: 0

; DB: 4 Gaps: 0

## US-09-763-153-1 (1-1245) x US-09-667-570A-3 (1-419)

QY 1 GCCAACTCCTCTCGAGGAGCTCCGTCACAGACGCTGGAGCGGAGTGCATAGAGGAG 60

DB 1 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 20

QY 61 ATCTGTGACTCGAGAGGCCAAGAAATTTCCAAATGTTGGATGACACACTGGCTTC 120

DB 21 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspThrLeuAlaPhe 40

QY 121 TGGTCCAAAGCAGCTCGACGGTGACAGTGTGCTTGTCTTGGCCCTGGAGACCCGCTGGCC 180

DB 41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60

QY 181 AGCTGTGCTGGGCGGAGCGGCTTCGCCAGCGGAGTGCATCGAGCTTCAGTGCAGTGC 240

DB 61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80

QY 241 CCACGGCTGGAGGCGGCTTCGCCAGCGGAGTGCATCGAGCTTCCTCAATTGCTCGCTG 300

DB 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100

QY 301 GACACGGCGGCTGACGCAATCTACTGCTTAGAGAGTGGCTGGCGGCTGTAGCTGT 360

DB 101 AspAsnGlyGlyCysThrHisThrCysLeuGluValGlyTrpArgArgCysSerCys 120

QY 361 GCGCTGCTGCTACAGCTGGGAGGAGCTTCCTGCTGCTGCTACCCCGCAGTGAAGTCCCT 420

DB 121 AlaProGlyTyrLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 140

QY 421 TGTGGAGGCGCTGAAGCGGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480

DB 141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160

QY 481 GACCAAGAGAACCAAGTAGATCCGCGGCTCATTTGATGGAAGATCACCAGCGGGAGAC 540

DB 161 AspGlnGluAspGlnValAspProArgLeuLeuLeuAspGlyLysMetThrArgArgGlyAsp 180

QY 541 AGCCCTGCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

DB 181 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 200

QY 601 ATCCACCCCTCTCTGGTGTGCTGACACGGCGGCGGCGGCGGCGGCGGCGGCGGCTCT 660

Db 201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 220  
QY 661 GTCAGGCTGGAGAGTATGACCTGGCGGCTGGGAGAAAGTGGAGCTGGACCTGGACATC 720  
Db 221 ValArgLeuGlyGluTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 240  
QY 721 AAGGAGGTCTCTCCACCCCAACTACAGCAAGAGCACCACCAATGACATCGGACATG 780  
Db 241 LysGluValPheValHisProAsnTyrSerLysSerThrAspAsnAspIleAlaLeu 260  
QY 781 CTCACCTGGCCCGCCGCCACCTCTCCAGACCATAGTCCCATCTGCTCCCGGAC 840  
Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProLecCysLeuProAsp 280  
QY 841 AGCGGCTTCAGAGCGGCGGCTCAATCAGCGGCGGCGGAGACCCCTCGTGGCGGCTCG 900  
Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
QY 901 GGCTACACAGCAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
Db 301 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
QY 961 ATCAAGATTCCTGCTGGCGCACATGAGTCGAGCGAGTGCATGAGCAACATGCTCTCT 1020  
Db 321 IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
QY 1021 GAGAACATGCTGTCGGGCGCATCTCCGCGGAGCGGCGGAGGAGGAGGAGGAGGAGG 1080  
Db 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
QY 1081 GGGGGCGGCGGCTGCTCTCCACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140  
Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
QY 1141 GGTGAGGCTGCTGGCTCTTCCAACTACAGGCGGTTTACACCAAGTTCAGCGGCTACCTC 1200  
Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValThrLysValSerArgTrpLeu 400  
QY 1201 GACTGTGATCATGGGCGACATCAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1245  
Db 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415

## RESULT 4

PCT-US92-10242-1

; Sequence 1, Application PC/TUS9210242

## ; GENERAL INFORMATION:

; APPLICANT: Griffin, John H.

; APPLICANT: Mesters, Rolf

; TITLE OF INVENTION: Serine Protease-Derived Polypeptides and

; TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods

; TITLE OF INVENTION: for Inhibiting Coagulation

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Office of Patent Counsel, The Scripps

; ADDRESSEE: Research Institute

; STREET: 10666 North Torrey Pines Road, TPC 8

; CITY: La Jolla

; STATE: CA

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; APPLICATION NUMBER: PCT/US92/10242

; FILING DATE: 19921118

; CLASSIFICATION:

; APPLICATION NUMBER: US 07/793,989

; FILING DATE: 18-NOV-1991

ATTORNEY/AGENT INFORMATION:  
 NAME: Fitting, Thomas  
 REGISTRATION NUMBER: 34,163  
 REFERENCE/DOCKET NUMBER: SCRO472P  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 619-554-2937  
 TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 419 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear

MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO

FEATURE:  
 NAME/KEY: Region  
 LOCATION: 1..157  
 OTHER INFORMATION: /note= "Protein C Light Chain"

FEATURE:  
 NAME/KEY: Region  
 LOCATION: 158..169  
 OTHER INFORMATION: /note= "Protein C Activation"  
 OTHER INFORMATION: Peptide"

FEATURE:  
 NAME/KEY: Region  
 LOCATION: 170...419  
 OTHER INFORMATION: /note= "Protein C Heavy Chain"

PCT-US92-10242-1  
 Alignment Scores:  
 Pred. No.: 3 34e-167 Length: 419  
 Score: 2298.00 Matches: 415  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.33% Indels: 0  
 DB: 5 Gaps: 0

US-09-763-153-1 (1-1245) x PCT-US92-10242-1 (1-419)

QY 1 GCCAATCTCTCTGAGAGCTCTCTACAGAGCGCTGGAGCGGAGTGCATAGAGGAG 60  
 Db 1 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 20  
 QY 61 ATCTGCTACTTCGAGGCGCCAGGAAATTTCCAAATCTGGATGACACACTGGCCCTC 120  
 Db 21 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 40  
 QY 121 TGTCTCAAGCACGTCGACGCTGACCACTGCTTGGCTTGTGCTTGGAGCACCCTGCGCC 180  
 Db 41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60  
 QY 181 AGCTGTGCTGCGGCGCACGAGTGCATCGAGCGCATCGGAGCTTCAGCTCGCACTGC 240  
 Db 61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
 QY 241 CGCAGCGGCTGGAGGCGCTCTCTCCAGCGGAGTGCATCTCTCAATTGCTCGCTG 300  
 Db 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
 QY 301 GACAAGCGGCTGTCACGATTAATCTGCTAGAGGAGTGGGCTGGGCGGCTGTAGCTGT 360  
 Db 101 AspAsnGlyCysThrHisThrCysLeuGluGluValGlyTrpArgArgCysSerCys 120  
 QY 361 GCGCTGCTGCTAACGCTGGGGAGACGCTCTGCTGCACTGTCACCCGCACTGAAGTTCCT 420  
 Db 121 AlaProGlyTrpLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 140  
 QY 421 TGTGGGAGGCGCTTGAAGCGGATGGAGAGACGCGAGTCACTGAACAGACACAGAA 480  
 Db 141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgPheThrGlu 160  
 QY 481 GACCAAGAGACCAAGTAGATCCGCGCTCATTTGATGGGAAGATGACACGCGGGGAGAC 540

Db 161 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 180  
 QY 541 AGCCCTCGGAGGTGGTCTCTGCTGACTCAAGAAGAGAGTGGCTGGCGGGAGTGCCTC 600  
 Db 181 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 200  
 QY 601 ATCCACCCCTCTCTGGTGTCTGACAGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 Db 201 IleHisProSerTrpValLeuThrAlaHisCysMetAspGluSerLysLysLeuLeu 220  
 QY 661 GTCAGGCTTGGAGAGTATGACCTGCGCGCTGGAGAGAGTGGAGAGTGGAGTGGAGTGC 720  
 Db 221 ValArgLeuGlyGluTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 240  
 QY 721 AAGGAGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
 Db 241 LysGluValPheValHisProAsnTrpSerLysSerThrThrAspAsnAspIleAlaLeu 260  
 QY 781 CTGCACCTGGCGCCAGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
 Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 280  
 QY 841 AGCGGCTTGCAGAGCGGAGCTCAATCAGCGCCGCGAGAGACCTCTGCTGCTGCTGCTG 900  
 Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
 QY 901 GGCTATCACAGCGCGAG 960  
 Db 301 GlyTrpHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
 QY 961 ATCAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
 Db 321 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
 QY 1021 GAGAACATGCTGTGTGCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
 Db 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
 QY 1081 GGGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
 Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
 QY 1141 GGTGAGGCTGTGGGCTCTCTCAACATACGCGCTTTACACCAAGTCCAGCGCTACCTC 1200  
 Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValTrpLysValSerArgTrpLeu 400  
 QY 1201 GACTGGATCCATGGGCGCATCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245  
 Db 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415

RESULT 5  
 US-08-756-506-2  
 ; Sequence 2, Application US/08756506  
 ; Patent No. 5905185  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Garner, Ian  
 ; APPLICANT: Cottingham, Ian R.  
 ; APPLICANT: Temperley, Simon M.  
 ; APPLICANT: Foster, Donald C.  
 ; APPLICANT: Sprecher, Cindy A.  
 ; APPLICANT: Prunkard, Donna E.  
 ; TITLE OF INVENTION: PROTEIN C PRODUCTION IN TRANSGENIC  
 ; NUMBER OF SEQUENCES: 25  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: ZymoGenetics, Inc.  
 ; STREET: 1201 Eastlake Avenue East  
 ; CITY: Seattle  
 ; STATE: WA  
 ; COUNTRY: USA  
 ; ZIP: 98102  
 ; COMPUTER READABLE FORM:

```

;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/756.506
;

```

Alignment Scores:	
Pred. No.:	3 42a-157
Score:	2298.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	98.93%
DB:	2
Length:	460
Matches:	415
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-763-153-1 (1-1245) x US-08-756-506-2 (1-460)

1	QY	1	GC	CAACTCCTCTCTGGAGGAGCTCGTACACGACACCCTGGAGCGGGAGTGCATAGAGGAG	60
43	Db	43	Ala	AsnSerPheLeuGluLeuLeuArgHisSerLeuGluArgGluCysIleLeuGlu	62
61	QY	61	ATC	TGTGACTCTGAGGAGGCCAAGAAATTTTCCAAAATGTGGATGACACACGTGGCGCTTC	120
63	Db	63	Ile	CysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe	82
121	QY	121	TGG	TCCAAGCACGTCGACGGGTGACCACTGCTTGCTTGCTTGGCCCTGGAGCACCCGCTGC	180
83	Db	83	Trp	SerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla	102
181	QY	181	AGC	CTGTCTCGGGCACGGCAGTCGCATCGACGGCATCGCAGCTTCAGCTCCGACTGC	240
103	Db	103	Ser	LeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys	122
241	QY	241	GC	CAGCGCTGGAGGGCGCGCTTCTGCCAGCGCAGGTGAGCTTCCCAATCTCGCTCG	300
123	Db	123	Arg	SerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu	142
301	QY	301	GAC	AACGGCGGCTGCACGATTA	360
143	Db	143	Asp	AsnGlyGlyCysThrHisIleCysLeuGluGluValGlyTrpArgArgCysSerCys	162
361	QY	361	CG	CTTGCTACAAAGTGGGGGACACCTCTCTGAGTGTACACCCCGAGTGAGTTCCT	420
163	Db	163	Ala	ProGlyTyrLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro	182
421	QY	421	TGT	GGGAGGCCCTGGAAAGCGGATGAGAAAGCGCAGTCACCTGAAACGACACAGAA	480
193	Db	193	Cys	GlyArgProIlePysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu	202
481	QY	481	GAC	CAAGAAGACCAAGTAGATCCGCGGCTTCAATGATGGGAAGATGACACAGCGGGGAG	540
203	Db	203	Asp	GlnLysAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp	222
541	QY	541	AG	CCCCCTGGCAGGTGCTCTGCTGACTCAAGAAGAAAGCTGGCTGCGGGGCGTGC	600
223	Db	223	Ser	ProIleGlnValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu	242

CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Sawislak, Deborah A  
REGISTRATION NUMBER: 37,438  
REFERENCE/DOCKET NUMBER: 95-28  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-442-6672  
TELEFAX: 206-442-6678  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 460 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
08-756-506-4

Alignment Scores:  
Pred. No.: 3,42e-167 Length: 460  
Score: 2298.00 Matches: 415  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.33% Indels: 0  
DB: 2 Gaps: 0

US-09-763-153-1 (1-1245) x US-08-756-506-4 (1-460)

```
QY 1 GCCAACTCCTCTCGAGGAGCTCCGTCACAGCACCTGGAGCGGGAGTGATAGAGGAG 60
Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62
QY 61 ATCTGTGACTTCGAGGAGGCCAAGGAAATTTTCCAAAATGTGGATGACACACTGGCCCTTC 120
Db 63 IleCysAspPheGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82
QY 121 TGGTCAACAGCGTCGAGGTCACAGCTGCTTGTGCTTGGCCCTGGAGACCCCGTGGCC 180
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102
QY 181 AGCCGTGTCTGGGGCAGCGACGTCATCGACGCGCATCGGCGCTTCAGCTGCAGCTGC 240
Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122
QY 241 CGCAGCGCTGGAGGGCGCTTCTCGCAGCGCGAGGTGAGCTTCTCAATTGCTCGCTG 300
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142
QY 301 GACACGCGGCTGCACCATCTACTGCTCTAGAGAGGTGGCGTGGCGCGCTGAGCTGT 360
Db 143 AspAsnGlyGlyCysThrHisTyrCysLeuGluGluValGlyTrpArgArgCysSerCys 162
QY 361 GCGCTGTGCTACAAGTGGGGCAGCAGCTCCTGAGTGTACCGCGCAGTGAAGTTCCT 420
Db 163 AlaProGlyTyrLysLeuGlyAspLeuLeuGlnCysHisProAlaValLysPhePro 182
QY 421 TGTGGAGGCCCTGGAACGGATGAGAGGAGCGAGTCACTCACTCAACACGACACAGAA 480
Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202
QY 481 GACCAAGAAGACCAAGTAGATCCGCGCTCATTTGATGGAGAGATGACCAGCGGGGAGAC 540
Db 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222
QY 541 AGCCCTCGCAGGTGCTCTGCTGCTCAAGAGAGCTGGCGTGGGGGCGAGTGCCTC 600
Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242
QY 601 ATCCACCCCTCTGCGGTGCTGACAGCGGCCCTGATGATGATGATGATGATGATGATGAT 660
Db 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 262
QY 661 GTCAGGCTTGGAGAGTATGACCTGGCGCTGGGAGAGTGGGAGCTGGAGCTGGACATC 720
Db 263 ValArgLeuGlyGlyLysLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 282
```

```
QY 721 AAGGAGGTCTTCGTCCACCCCACTACAGCAAGAGCACCCAGCAATGATGATGACTG 780
Db 283 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 302
QY 781 CTGCACTGGCGCCAGCGCCACCTCTCGCAGACCATAGTGCCTGCTCCCTCCCGGAC 840
Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322
QY 841 AGCGGCTTCGAGAGCGCGAGCTCAATCAGCGCGCCAGGAGACCTCTGTCGACGGCTG 900
Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342
QY 901 GGCTACACACAGCCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
Db 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362
QY 961 ATCAAGATTCCCGTGGTCCGCGACAATGAGTGCAGCGAGGTGCATGAGCAACATGGTCT 1020
Db 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382
QY 1021 GAGAATCATGCTGTGCGGCGCATCTCGGGGACCGGAGGAGTGCCTCGAGGGCGACAGT 1080
Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402
QY 1081 GGGGGGCGCATGTCGCTCTCCACGCGACCTGCTTCTGTTGGGCGCTGGTGGAGCTGG 1140
Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422
QY 1141 GGTGAGGCTGTGGGCTCTTCACTACAGTACGGGCTTTACACCAAGTCACCGCTACCTC 1200
Db 423 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 442
QY 1201 GACTGTGATCCATGGCGCATCATCAGACAAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245
Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457
```

## RESULT 7

5225537-2  
; Patent No. 5225537  
; APPLICANT: FOSTER, DONALD  
; TITLE OF INVENTION: METHODS FOR PRODUCING HYBRID  
; PHOSPHOLIPID-BINDING PROTEINS  
; NUMBER OF SEQUENCES: 14  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/459,082  
; FILING DATE: 29-DEC-1989  
; SEQ ID NO: 2  
; LENGTH: 461  
5225537-2

Alignment Scores:  
Pred. No.: 3,42e-167 Length: 461  
Score: 2298.00 Matches: 415  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.33% Indels: 0  
DB: 2 Gaps: 0

US-09-763-153-1 (1-1245) x 5225537-2 (1-461)

```
QY 1 GCCAACTCCTCTCGAGGAGCTCCGTCACAGCACCTGGAGCGGGAGTGATAGAGGAG 60
Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62
QY 61 ATCTGTGACTTCGAGGAGGCCAAGGAAATTTTCCAAAATGTGGATGACACACTGGCGCTTC 120
Db 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82
QY 121 TGTCAACAGCGCTGCAGGTCACAGTGTGCTTGTGCTTGGCCCTGGAGACCCCGTGGCC 180
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102
```

QY 181 AGCTGTGCTCGGGCAGCGGCATGATGATGACGGCATCGGAGCTTCAGCTTCAGCTGCGACTGC 240  
Db 103 SerLeuCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
QY 241 CGCAGCGCTGGAGGGCGCTTCGCCAGCGCAGCTGAGCTTCCTCAATGCTCGCTG 300  
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
QY 301 GACAAAGCGCGCTGCACCATCTACTGCTAGAGAGGTGGCTGGCGCGCTGAGCTGT 360  
Db 143 AspAsnGlyGlyCysThrHisThrCysLeuGluValGlyTrpArgArgCysSerCys 162  
QY 361 GCGCTGCTGCTACAGCTGGGGAGCAGCTCTGCTGAGTGTACCCCGCAGTGAAGTTCCT 420  
Db 163 AlaProGlyTrpLysLeuGlyAspLeuGlyHisProAlaValLysPhePro 182  
QY 421 TGTGGAGGCGCTGGAAGCGGATGGAGAGAGCGCAGTCACTGAAACGAGACACAGAA 480  
Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
QY 481 GACCAAGAGACCAAGTAGATCCCGGCTCATTTGATGGAGAGATGACCGCGGGGAGAC 540  
Db 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
QY 541 AGCCCTGCGCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242  
QY 601 ATCCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
Db 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu 262  
QY 661 GTCAGGCTTGGAGATGATGCTGCGGCGCTGGGAGAGTGGAGCTGGAGCTGGACATC 720  
Db 263 ValArgLeuGlyGlyTrpAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
QY 721 AAGAGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
Db 283 LysGluValPheValHisProAsnTrpSerLysSerThrAspAsnAspIleAlaLeu 302  
QY 781 CTGACCTTGGCGCCAGCGCCTCTGCGAGACCATAGTGCATCTGCTGCTGCTGCTGCT 840  
Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
QY 841 AGCGCTTGGCAGCGGAGCTCAATCAGCGCGCGCAGGAGACCCCTGCTGACGGGCTGG 900  
Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
QY 901 GCTACACAGCGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
Db 343 GlyTrpHisSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
QY 961 ATCAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
Db 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
QY 1021 GAGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
QY 1081 GGGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
QY 1141 GGTGAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
Db 423 GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValThrLysValSerArgTrpLeu 442  
QY 1201 GACTGGATTCATGGGCATCATGAGACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245  
Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457  
RESULT 8

5460953-3  
; Patent No. 5460953  
; APPLICANT: GERLITZ, BRUCE E.; GRINNELL, BRIAN W.  
; TITLE OF INVENTION: VECTORS AND COMPOUNDS FOR EXPRESSION OF  
; GLYCOSYLATION MUTANTS OF HUMAN PROTEIN C  
; NUMBER OF SEQUENCES: 3  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/93,217  
; FILING DATE: 09-SEP-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 628,063  
; FILING DATE: 21-DEC-1990  
; APPLICATION NUMBER: 484,081  
; FILING DATE: 23-FEB-1990  
; SEQ ID NO: 3  
; LENGTH: 461  
5460953-3  
Alignment Scores:  
Pred. No.: 9,78e-167 Length: 461  
Score: 2292.00 Matches: 414  
Percent Similarity: 99.76% Conservative: 0  
Best Local Similarity: 99.76% Mismatches: 1  
Query Match: 98.07% Indels: 0  
DB: 6 Gaps: 0  
US-09-763-153-1 (1-1245) x 5460953-3 (1-461)  
QY 1 GCCAATCTCTCTCTGAGGAGCTCGCTCAGCAGCTGGAGCGGAGTGATAGAGGAG 60  
Db 43 AlaAsnSerPheLeuGluLeuArgHisSerLeuGluArgGluCysIleGluGlu 62  
QY 61 ATCTGTGACTTTCGAGGAGCGCAAGGAAATTTTCCAAATGTGGATGACACTGGCCTTC 120  
Db 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
QY 121 TGTCTCAAGACACTGACAGGTGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
QY 181 AGCTGTGCTGCGGCGCAGCGCAGCTGCATCGAGCATCGGAGCTTCAGCTTCAGCTGCG 240  
Db 103 SerLeuCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
QY 241 CGCAGCGCTGGAGGGCGCTTCTGCCACCGCGAGGTGAGCTTCCTCAATGCTGCTGCTG 300  
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
QY 301 GACAAGCGGCGCTGCACCATCTACTGCTAGAGAGGTGGCTGGCGCGCTGCTGCTGCT 360  
Db 143 AspAsnGlyGlyCysThrHisThrCysLeuGluGluValGlyTrpArgArgCysSerCys 162  
QY 361 GCGCTGCTGCTACAAAGCTGGGGAGCAGCTCCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
Db 163 AlaProGlyTrpLysLeuGlyAspLeuLeuGlnCysHisProAlaValLysPhePro 182  
QY 421 TGTGGAGGCGCTTGGAGCGGATGGAGAGACCGCAGTCACTGAAACGAGACACAGAA 480  
Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
QY 481 GACCAAGAGACCAAGTAGATCCCGGCTCATTTGATGGAGAGATGACCGCGGGGAGAC 540  
Db 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
QY 541 AGCCCTGCGCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242  
QY 601 ATCCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
Db 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu 262  
QY 661 GTCAGGCTTGGAGATGATGCTGCGGCGCTGGGAGAGTGGAGCTGGAGCTGGACATC 720



Db 263 ValArgLeuGlyGluTyArgAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
QY 721 AAGGAGGCTTCGTCACCCCAACACACAGAGACACACGCAATGACATCGACTG 780  
Db 283 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 302  
QY 781 CTCGACCTGGCCAGCCGACCTCTCGCAGACCATAGTCCCATCTGCTCCCGGAC 840  
Db 303 LeuHisLeuAlaGluProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
QY 841 ACGGCTTGGACAGCGGAGCTCAATCAGCCCGCCAGGACGCTCGTGACGGCTGG 900  
Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
QY 901 GCTACACAGCAGCGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
Db 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
QY 961 ATCAAGATTCGCTGCTCCGACCAATGAGTGCAGCGAGGTCAATGAGCAACATGTTGT 1020  
Db 363 IleLysIleProValValProHisGlnGluCysSerGluValMetSerAsnMetValSer 382  
QY 1021 GAGACATGCTGTGCGGGCATCTCGGGACCGGACGAGGATGCTGCGAGGGCGACAGT 1080  
Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
QY 1081 GGGGGCCCATGGTCCCTTCCACACATGAGTGCAGCGAGGTCAATGAGCAACATGTTGT 1140  
Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
QY 1141 GGTGAGGCTGTGGCTCTTCAACATPACGGCGTTTACACAAAGTCAGCGCTACCTC 1200  
Db 423 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 442  
QY 1201 GACTGGATCCATGGCAGATCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245  
Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

## RESULT 9

5270178-17

; Patent No. 5270178

; APPLICANT: GERLITZ, BRUCE E.; GRINNELL, BRIAN W.

; TITLE OF INVENTION: VECTORS AND COMPOUNDS FOR EXPRESSION OF

; ZYMOGEN FORMS OF HUMAN PROTEIN C

; NUMBER OF SEQUENCES: 21

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/484,133

; FILING DATE: 23-FEB-1990

; SEQ ID NO:17:

LENGTH: 461

5270178-17

## Alignment Scores:

Pred. No.:	9, 29e-165	Length:	461
Score:	2266.00	Matches:	409
Percent Similarity:	99.28%	Conservative:	3
Best Local Similarity:	98.55%	Mismatches:	0
Query Match:	96.96%	Indels:	0
DB:	6	Gaps:	0

US-09-763-153-1 (1-1245) x 5270178-17 (1-461)

QY 1 GCCAACTCCCTTCTGGAGGAGCTCCGTCACAGCAGCGCTGGAGCGGGAGTGCATAGAGGAG 60  
Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
QY 61 ATCTGTGACTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
Db 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
QY 121 TGGTCCAAAGCAGCTGCGAGGTGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
QY 181 AGCCTGTGCTGCGGGCACGCGCATGTCATCGACGGCATCGCAGCTTCACTGCTGACTGC 240  
Db 103 SerLeucCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
QY 241 CGCAGCGCTGGGAGGCGCCCTCTGCGCAGCGGAGGTGAGCTTCCTCAATGCTCGCTG 300  
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
QY 301 GACAAACGCGCTGACGCGCATTTACTGCTAGAGAGGTGGCTGGCGGCGCTTAGCTGT 360  
Db 143 AspAsnGlyGlyCysThrHisTyrCysLeuGluGluValGlyTrpArgArgCysSerCys 162  
QY 361 GCGCTGGCTACAGCTGGGGGACGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
Db 163 AlaProGlyTyrLysLeuGlyAspLeuLeuGlnCysHisProAlaValLysPhePro 182  
QY 421 TGTGGGAGCGCTGGAGCGGATGGAGAGACGAGCAAGCGAGTCACCTGAAACGAGACAGAA 480  
Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
QY 481 GACCAAGAACACCAAGTAGATCCGCGCTCATTGATGGGAGATGACACGAGCGGGAGAC 540  
Db 203 AspGlnGluAspGlnValPheProArgLeuIleAsnGlyLysMetThrArgArgGlyAsp 222  
QY 541 AGCCCTGGCAGGTGCTGCTGGACTCAAAGAAGAAGCTGCCTGCGGGCGAGTGCCTC 600  
Db 223 SerProTrpGlnValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242  
QY 601 ATCCACCCCTCCCTGGTGTGACAGCGCGCTGATGATGATGATGATGATGATGATGATG 660  
Db 243 IleHisProSerTrpValLeuThrAlaHisCysMetAspGluSerLysLysLeuIle 262  
QY 661 GTCAGCTTGGAGAGTATGACCTGCGCGCTGGGAGAGTGGGAGCTGACCTGGACATC 720  
Db 263 ValArgIleGlyLysTrpAspLeuArgArgTrpGluLysTrpLeuLeuAspLeuAspIle 282  
QY 721 AAGGAGGTCTTCTGCTCCACCCCACTACAGCAGAGAGCACCAGCAATGACATCGCAGT 780  
Db 283 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 302  
QY 781 CTGACCTGCGCCAGCGCCGCTCTCGCAGACCATAGTGCCTGCTGCTGCTGCTGCTGCTG 840  
Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
QY 841 ACGGCTTGGCAGCGCGAGCTCAATCAGCGCGCGGACGAGAGACCTCTGCTGACGGCTG 900  
Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
QY 901 GGTACCCAGCAGCGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
Db 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
QY 961 ATCAAGATTCGCTGCTCCGACCAATGAGTGCAGCAGGAGTGCATGAGCAACATGTTGT 1020  
Db 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
QY 1021 GAGACATGCTGTGCGGGCATCTCGGGGACCGGACGAGGATGCTGCGGAGGCGGACAGT 1080  
Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
QY 1081 GGGGGCCCATGCTGCTCCCTTCCAGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
QY 1141 GGTGAGGCTGTGGCTCTTCCAACTACGCGCTTTACACCAAGTCAAGCTGCTGCTGCT 1200  
Db 423 GlyGluGlyCysGlyLeuLeuHisAsnValGlyValTyrThrHisValSerArgTyrLeu 442  
QY 1201 GACTGGATCCATGGCAGATCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245  
Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

RESULT 10  
5270178-18  
; Patent No. 5270178  
; APPLICANT: GERLITZ, BRUCE E.; GRINNELL, BRIAN W.  
; TITLE OF INVENTION: VECTORS AND COMPOUNDS FOR EXPRESSION OF  
; ZYMOGEN FORMS OF HUMAN PROTEIN C  
; NUMBER OF SEQUENCES: 21  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/484,133  
; FILING DATE: 23-FEB-1990  
; SEQ ID NO:18:  
; LENGTH: 461  
5270178-18  
Alignment Scores:  
Pred. No.: 9,29e-165 Length: 461  
Score: 2266.00 Matches: 409  
Percent Similarity: 99.28% Conservativity: 3  
Best Local Similarity: 98.55% Mismatches: 3  
Query Match: 96.96% Indels: 0  
Gaps: 0  
09-763-153-1 (1-1245) x 5270178-18 (1-461)  
QY 1 GCCAACTCTTCTGAGGAGCTCCGTCACAGCAGCCCTGGAGCGGAGTGCATAGAGGAG 60  
Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
QY 61 ATCTGTGACTTCGAGAGGCCAAGGAATTTTCCAAATGTGGATGACACACTGGCCCTC 120  
Db 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
QY 121 TGGTCCAAAGCAGTGCAGCGGTGACAGTGTGGTCTTGGCTTGGAGTGGAGTGGAGTGG 180  
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
QY 181 AGCCTGTGCTGGGCGCAGCGGACGAGTGCATCGACGCGTTCAGTGCAGTGCAGTGC 240  
Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
QY 241 CGCAGCGGCTGGAGGCGGCTCTGCCAGCGGAGGTGAGTCTCAATTCGCTCGCTG 300  
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
QY 301 GACAACGGCGGCTGCACGATTAATCTGCTAGAGAGTGGGTGGCGGCTGTAGCTGT 360  
Db 143 AspAsnGlyGlyCysThrHisTyrcysLeuGluValGlyTrpArgArgCysSerCys 162  
QY 361 GCGCTGTGCTACAGTGGGCGGACGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
Db 163 AlaProGlyTyrllysLeuGlyAspAspLeuLeuGlnCysHisProAlaVallysPhePro 182  
QY 421 TGTGGAGGCGCTCGAAGCGGATGGAGAGAGAGCGAGTCACTGAAACGACACAGAA 480  
Db 183 CysGlyArgProTrpIysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
QY 481 GACCAAGAACAGCAAGTAGATCCGGCTCATTTGATGGGAAGATGACAGCGGCGGAGAC 540  
Db 203 AspGlnGluAspGlnValPheProArgLeuLeuAsnGlyLysMetThrArgArgGlyAsp 222  
QY 541 AGCCCTGGCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 242  
QY 601 ATCCACCCCTCTGGGTGTGACAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
Db 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGlnSerLysLysLeuIle 262  
QY 661 GTCAGGCTTGGAGTATGACCTGGCGGCTGGAGAGTGGAGTGGAGTGGAGTGGAGTGG 720  
Db 263 ValArgIleGlyGluTyrlAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 282

QY 721 AAGGAGTCTTCTGCTCCACCCCACTACAGCAAGAGACACACCGACCAATGACATCGACTG 780  
Db 283 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 302  
QY 781 CTGCACTGGCCAGCGCCGCCACCTCTCGCAGACCATAGTGCCTCCCTCCCGGAC 840  
Db 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
QY 841 AGCGGCTTTCAGAGCGCGAGCTCAATCAGCGCGCGCAGGAGACCTCTGACGGGCTG 900  
Db 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
QY 901 GGCTACACAGCAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
Db 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
QY 961 ATCAAGATTCCTGCTGCTCCCGCAATGAGTGCAGCGAGTGCATGAGCAACATGGTGTCT 1020  
Db 363 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 382  
QY 1021 GAGAACATGCTGTGCGGCGCATCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080  
Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
QY 1081 GGGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
QY 1141 GGTGAGGCTGTGGGCTCTTCAACTACGCGGCTTTACACCAAGTTCAGCCGCTACCTC 1200  
Db 423 GlyGlyGlyCysGlyLeuLeuHisAsnValGlyValTyrThrHisValSerArgTyrLeu 442  
QY 1201 GACTGTGATCCATGGCGCATCAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245  
Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457  
RESULT 11  
5270178-13  
; Patent No. 5270178  
; APPLICANT: GERLITZ, BRUCE E.; GRINNELL, BRIAN W.  
; TITLE OF INVENTION: VECTORS AND COMPOUNDS FOR EXPRESSION OF  
; ZYMOGEN FORMS OF HUMAN PROTEIN C  
; NUMBER OF SEQUENCES: 21  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/484,133  
; FILING DATE: 23-FEB-1990  
; SEQ ID NO:13:  
; LENGTH: 460  
5270178-13  
Alignment Scores:  
Pred. No.: 2,43e-164 Length: 460  
Score: 2260.50 Matches: 409  
Percent Similarity: 99.28% Conservativity: 3  
Best Local Similarity: 98.55% Mismatches: 2  
Query Match: 96.73% Indels: 1  
Gaps: 1  
US-09-763-153-1 (1-1245) x 5270178-13 (1-460)  
QY 1 GCCAACTCTTCTGAGGAGCTCCGTCACAGCAGCCCTGGAGCGGAGTGCATAGAGAG 60  
Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
QY 61 ATCTGTGACTTCGAGAGGCCAAGGAATTTTCCAAATGTGGATGACACACTGGCCCTC 120  
Db 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
QY 121 TGGTCCAAAGCAGTGCAGCGGTGACAGTGTGGTCTTGGCTTGGAGTGGAGTGGAGTGG 180  
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
QY 181 AGCCTGTGCTGGGCGCAGCGGACGAGTGCATCGACGCGTTCAGTGCAGTGCAGTGC 240

Db 103 SerLeuCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
QY 241 CCAGCGGCTGGAGGGCGCTTCGCGAGCGGAGGTGAGCTTCTCAATTGCTCGCTG 300  
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
QY 301 GACAACGGCGGCTGACCACTTACCTAGAGGAGGTGGCTGGCGGCTGTAGCTGT 360  
Db 143 AspAsnGlyGlyCysThrHisGlyCysLeuGluValGlyTrpArgCysSerCys 162  
QY 361 GCGCTGGCTACAGCTGGGAGAGCACTCTCGAGTGTCTACCGCGAGTGAAGTTCCT 420  
Db 163 AlaProGlyTyrLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro 182  
QY 421 TGTGGAGGCGCTGGAACCGGATGGAGAGAGCGCAGTCACTGAAACGAGACAGAA 480  
Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
QY 481 GACCAAGAGACCAAGTAGATCCCGGCTATTGATGGAGAGATGACAGCGGGGAGAC 540  
Db 203 AspGlnGluAspGlnValAspProArgLeu---AsnGlyLysMetThrArgArgGlyAsp 221  
QY 541 AGCCCTGGCAGGTGCTCTGCTGACTCAAGAGAGGCTGCGCTCGGGGCGAGTGTCT 600  
Db 222 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 241  
QY 601 ATCCACCCCTCTCGGTGCTGACAGCGCCCACTGCGATGATGATGATGATGATGAT 660  
Db 242 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu 261  
QY 661 CTCAGGCTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
Db 262 ValArgIleGlyGluTyrAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAsp 281  
QY 721 AAGAGGTCTCTGCTCCACCCCACTACAGAGAGCACCACCAATGATGATGATGATGAT 780  
Db 282 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 301  
QY 781 CTGACCTGGCGCCAGCGGAGCTCAATCAGCGCGGCGGAGAGACCTCTGTCAGCGGCTG 360  
Db 302 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 321  
QY 841 AGCGGCTTCAGAGCGGAGCTCAATCAGCGCGGCGGAGAGACCTCTGTCAGCGGCTG 900  
Db 322 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 341  
QY 901 GGCTTACCACAGCGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
Db 342 GlyTyrHisSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 361  
QY 961 ATCAAGATTCCTGCTCCCGCAATAGTGCAGCGAGGTGATGAGCAACATGCTGTCT 1020  
Db 362 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 381  
QY 1021 GAGAATGCTGTGCGGCGATCTCTCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
Db 382 GluAsnMetLeuLysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 401  
QY 1081 GGGGGCCCATGCTGCTTCCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
Db 402 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 421  
QY 1141 GGTGAGGCTGTGGCTCTCTCAACTACGGCTTTACACCAAAAGTGGAGGAGGAGGAGGAG 1200  
Db 422 GlyGluGlyCysGlyLeuLeuHisAsnValGlyValTyrHisValSerArgTyrLeu 441  
QY 1201 GACTGATCATGGGCACATCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245  
Db 442 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 456  
RESULT 12  
5270178-14

; Patent No. 5270178  
; APPLICANT: BRUCE E. GRINNELL, BRIAN W.  
; TITLE OF INVENTION: VECTORS AND COMPOUNDS FOR EXPRESSION OF  
; ZYMOGEN FORMS OF HUMAN PROTEIN C  
; NUMBER OF SEQUENCES: 21  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/484,133  
; FILING DATE: 23-FEB-1990  
; SEQ ID NO: 14:  
; LENGTH: 460  
5270178-14  
Alignment Scores:  
Pred. No.: 2,43e-164 Length: 460  
Score: 2260.50 Matches: 409  
Percent Similarity: 99.28% Conservative: 3  
Best Local Similarity: 98.55% Mismatches: 2  
Query Match: 96.73% Indels: 1  
DB: 6 Gaps: 1  
US-09-763-153-1 (1-1245) x 5270178-14 (1-460)

QY 1 GCCAACTCTCTCTGGAGGAGCTCCCTCACAGCAGCTGGAGCGGAGTGCATAGAGAG 60  
Db 43 AlaAsnSerPheLeuGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
QY 61 ATCTGTGACTTCGAGGAGGCCAAGAAATTTTCCAAAATGTGGATGACACACTGGCCTTC 120  
Db 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
QY 121 TGGTCAACAGCAGCTCAGCGGTGACCACTGCTTGGCTTCCCTTGGAGCACCCGTCGCC 180  
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
QY 181 AGCCTGTGCTGGGGCAGCGCAGCTGCATCAGCGATCGGAGCTTACGCTGCGACTGC 240  
Db 103 SerLeuLysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
QY 241 CGCAGCGGCTGGAGGCGGCTTCTGCCAGCGGAGGTGAGCTTCTCAATTGCTCGCTG 300  
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
QY 301 GACAACGGCGGCTGACCACTTACTGCTAGAGAGGTGGCTGGCGGCTGTAGCTGT 360  
Db 143 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 162  
QY 361 GCGCTGTGCTACAAGCTGGGGAGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
Db 163 AlaProGlyTyrLysLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro 182  
QY 421 TGTGGAGGCGCTGGAAGCGGATGGAGAGAGCGCAGTCACTGAAACGAGACAGAA 480  
Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
QY 481 GACCAAGAGACCAAGTAGATCCCGGCTCATTTGATGGAGAGATGACAGCGGGGAGAC 540  
Db 203 AspGlnGluAspGlnValAspProArgLeu---AsnGlyLysMetThrArgArgGlyAsp 221  
QY 541 AGCCCTGGCAGGTGCTCTCTGCTGACTCAAGAGAGGCTGGCTGGCGGCGAGTGTCT 600  
Db 222 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 241  
QY 601 ATCCACCCCTCTGGGTGCTGACAGCGGCCCACTGCGATGATGATGATGATGATGATGAT 660  
Db 242 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu 261  
QY 661 GTCAGGCTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
Db 262 ValArgIleGlyGluTyrAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAsp 281  
QY 721 AAGAGGTCTCTGCTCCACCCCACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
Db 282 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 301

781	CTGCACCTGGCCAGCGCGCCACCTCTCCGACACCATAGTGCCCATCTCCGCGGAC	840
Qy		
302	LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleLeuProAsp	321
Db		
841	AGCGGCTTGCAGAGCGCGAGCTCAATCAGCGCGCCAGGAGACCTCGTGAAGGCTGG	900
Qy		
322	SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp	341
Db		
901	GGCTACACACAGCCGAGAGAGAGCGCCCAAGAGAAACCGCACTTCGTCCTCAACTTC	960
Qy		
342	GlyTyrHisSerSerArgGluGlyGluAlaLysArgAsnArgThrPheValLeuAsnPhe	361
Db		
961	ATCAAGATCCCGTGGTCCCGACAAATGAGTCAGCAGGAGTCATGAGCAACATGGTGCT	1020
Qy		
362	IleTyIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer	381
Db		
1021	GAGACATGCTGTGGCGGATCCTCGGGACCGGAGGATCCTCGAGGCGGACAGT	1080
Qy		
382	GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer	401
Db		
1081	GGGGGGCCATGGTGCTCCTCCACGGCACCTGGTTCTCGTGGGCCCTGGTGAAGTGG	1140
Qy		
402	GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp	421
Db		
1141	GGTGAGGCGTGGGCTCCTTCAACAATCAGCGGCTTTACACCAAGTCAGCCGCTACCTC	1200
Qy		
422	GlyGluGlyCysGlyLeuHisAsnValGlyValIleThrHisValSerArgTyrLeu	441
Db		
1201	GACTGGATCCATGGGCACATCAGAGACAAGGAGCCCCCAGAG	1245
Qy		
442	AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys	456
Db		

## RESULT 13

```

RESULT 13
US-09-065-872-1
; Sequence 1, Application US/09065872
; Patent No. 6162629
; GENERAL INFORMATION:
; APPLICANT: Baker, Jeffrey C
; APPLICANT: Carlson, Andrew D
; APPLICANT: Huang, Lihua
; APPLICANT: Sheliga, Theodore A
; TITLE OF INVENTION: Improved Methods for Processing Activated Protein C
; FILE REFERENCE: apc process patent
; CURRENT APPLICATION NUMBER: US/09/065,872
; CURRENT FILING DATE: 1998-04-24
; EARLIER APPLICATION NUMBER: 60/045,255
; EARLIER FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0

```

Alignment Scores:		
Pred. No.:	6.2e-164	Length:
Score:	2255.00	Matches:
Percent Similarity:	100.00%	Conservative:
Best Local Similarity:	100.00%	Mismatches:
Query Match:	96.49%	Indels:
DB:	4	Gaps:
		410
		406

US-09-763-153-1 (1-1245) X US-09-065-872-1 (1-410)

QY	28	CACAGCAGCCTGGAGCGGGAGTGCATAGAGGAGATCTGTGATTCGAGGAGGCGCAAGGAA	87
DB	1	HisSerLeuGluAruGluCysIleGluGluIleCysAspPheGluGluAAlaLysGlu	20
QY	88	ATTTTCCAAAATGTGGATGACACACATGGCGCTTCGGTCCCAAGCAGCTGCAGCGGTGACCAG	147
DB	21	TlePheGlnAsnValAsnAspTrileuAlaPheTrpSerLysHisValAspGluVaseGlu	40

**QY** 1228 AAGGAAGCCCCCAGAAG 1245  
| | | | | | | | | | | | | | |  
**Db** 401 LysGluAlaProGlnLys 406

**RESULT 14**

US-09-667-570A-1  
; Sequence 1, Application US/09667570A  
; Patent No. 6436397  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Jeffrey C  
; APPLICANT: Carlson, Andrew D  
; APPLICANT: Huang, Lihua  
; APPLICANT: Shelliga, Theodore A  
TITLE OF INVENTION: Improved Methods for Processing Activated Protein C  
FILE REFERENCE: X-11796A  
CURRENT APPLICATION NUMBER: US/09/667,570A  
CURRENT FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: 60/045,255  
PRIOR FILING DATE: 1997-04-28  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 1  
LENGTH: 410  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-667-570A-1

Alignment Scores:

Pred. No.:	Score:	Length:
28	6.2e-164	410
1	2255.00	Matches: 406
	100.00%	Conservative: 0
	Best Local Similarity: 100.00%	Mismatches: 0
	Query Match: 96.49%	Indels: 0
DB:	4	Gaps: 0

US-09-763-153-1 (1-1245) x US-09-667-570A-1 (1-410)

QY	28	CACAGAGCGCTGGAGCGGAGTGATGATGAGGAGATCTGTGACTTCGAGGAGGCCAAGGAA	87
Db	1	HisSerSerLeuGluArgLucylleGluIleCysAspPheGluAlaLysGlu	20
QY	88	ATTPTTCCAAAATGTGTATCACACTGGCTTTCTGGTCCAACGACGTGCAGCGTGACCAG	147
Db	21	IlePheGlnAsnValAspAspThrLeuAlaPheTrpSerLysHisValAspIlyAspGln	40
QY	148	TGCTTGGCTTGTCCCTTGGAGCACCCGTCGCCACGCTGTGTCTGGCGGCACGCGACGTGC	207
Db	41	CysLeuValLeuProLeuGluHisProCysAlaSerLeuCysCysglyHisclgYThrCys	60
QY	208	ATCGAGCGCATCGGACGCTTCAGCTCGGACTGCCGACGGCTGGGAGGGCGGCTTCCTGCG	267
Db	61	IleAspGlylleGlySerPheSerCysAspCysArgSerGlyTrpGluGlyArgPheCys	80
QY	268	CAGCGCGAGGTGAGCTTCCTCAATTGCTCGCTGGACACGCGCGGCTGCACGCATTACTGCG	327
Db	81	GlnArgGluValSerPheLeuAsnCysSerLeuAspAsnGlyCysThrHisTrCys	100
QY	328	CTAGAGAGGTGGCTGGCGGGCGCTGTACTGTGGCCCTGGCTACAAGCTGGGGGACGAC	387
Db	101	LeuGluGluValGlyTrpArgArgCysSerCysAlaProGlyTyTrLysLeuGlyAspasp	120
QY	388	CTCCTGCAGTGTCAACCCGACGTGAAGTTCCCTTGTGGGAGGCCCTGGAAAGCGGATGGNG	447
Db	121	LeuLeuGlnCysHisProAlaValLysPheProCysGlyArgproTrpLysargMetGlu	140
QY	448	AAGAAGCGCATGCACCTGAAACGAGACACAGAAAGACCAAGAAAGACCAAGTAGATCCGCGG	507
Db	141	LysLysArgSerHisLeuLysArgAspThrGluAspGlnGluAspGlnValAspProArg	160
QY	508	CTCATTTGATGGGAAGATGACACGAGGGGGGAGACAGCCCCTGGCAGGTGGTCTCTGTGGAC	567
Db	161	LeuIleAspGlyLysMetThrArgArgGlyAspSerProTrpGlnValValLeuLeuAsp	180

Db 42 AlaAsnSerPheLeuGluGluLeuArgHisSerLeuGluArgGluCysIleGluGlu 61  
QY 61 ATCTGTGACTTCGAGGAGGCAAGAAATTTCCAAAATGGAATGACACACACTGGCCCTC 120  
Db 62 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspThrLeuAlaPhe 81  
QY 121 TGGTCCAAAGCAGCTCGAGCGGTGACAGTGTGGTCTGTCTGGCTGGAGCACCCGTGGCC 180  
Db 82 ThrSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGlyHisProCysAla 101  
QY 181 AGCTGTGCTGGGCGCAGCGACGTGCATCGACGCGCAGGTGAGCTTCCATTTGCTCGCTG 240  
Db 102 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 121  
QY 241 CGCAGCGCTGGAGGCGCTCTTCCAGCGCGAGGTGAGCTTCCATTTGCTCGCTG 300  
Db 122 ArgSerGlyTyrPgluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 141  
QY 301 GACAAAGCGGCTGACGATTAATCTGCTAGAGGAGTGGCTGGCGCGCTGTAGCTGT 360  
142 AspAsnGlyGlyCysThrHisTyrCysLeuGluGluValGlyTyrArgArgCysSerCys 161  
361 GGGCTGCTACAGCTGGGCGAGACCTCTCTGAGTGTACCCCGCAGTGAAGTTCCT 420  
Db 162 AlaProGlyTyrLysLeuGlyAspLeuLeuGlnCysHisProAlaValLysPhePro 181  
QY 421 TGTGGAGCGCTGGAAGCGGTGAGAGAGCGCAGTCCACTGAAAGAGACACAGAA 480  
Db 182 CysGlyArgProTyrPgluGlyArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 201  
QY 481 GACCAAGAGAACCAAGTAGATCCCGGCTCATTTGATGGGAAGATGACCGCGGGAGAC 540  
Db 202 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 221  
QY 541 AGCCCTGGCAGGTGCTGCTGACTGCAAGAAAGAGTGGCTGGCGGGGAGTGC 600  
Db 222 SerProTyrGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 241  
QY 601 ATCCACCTCTCTGGTGTGCTGACAGCGCGCTGATGATGATGATGATGATGATGATGAT 660  
Db 242 IleHisProSerTyrPgluGlyArgPheAlaHisCysMetAspGluSerLysLeuLeu 261  
QY 661 GTGAGCTGGAGAGTATGACCTCGCGCGCTGGAGAGTGGAGCTGGAGCTGGAGATC 720  
Db 262 ValArgLeuGlyGlyTyrAspLeuArgArgArgGluLysTyrGluLeuAspLeuAspIle 281  
QY 721 AAGGAGTCTTCGTCACCCCACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
Db 282 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 301  
QY 781 CTGACCTGGCG 840  
302 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 321  
841 AGCGGCTTGGAGCGCGAGCTCAATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
322 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTyr 341  
QY 901 GGCTACACAGCGCGAG 960  
Db 342 GlyTyrHisSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 361  
QY 961 ATCAGATCTCCGTTGCCACAAAGTAGTGCAGAGGAGTGCAGAGGAGTGCAGAGGAGT 1020  
362 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 381  
QY 1021 GAGAACTGCTGTGCGGGCATCTCGGGGAGCGGAGGAGTGCCTGCGAGGGGAGAGT 1080  
Db 382 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 401  
QY 1081 GGGGGGCCCATGGTGGCTCTCCACGGCACCTGGTTCCTGGTGGCGCTGGTGGAGCTGG 1140

Db 402 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 421  
QY 1141 GGTGAGGCG---TGTGGCTCCTTCAAACTACGGGGTTTACACAAAGTTCAGCGCTAC 1197  
Db 422 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyr 441  
QY 1198 CTCGACTGGATCATCGGCACATCAGAGACAAGGAAGCCCCCAGAG 1245  
Db 442 LeuAspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457  
RESULT 16  
US-09-065-872-2  
; Sequence 2, Application US/09065872  
; Patent No. 6162629  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Jeffrey C  
; APPLICANT: Carlson, Andrew D  
; APPLICANT: Huang, Lihua  
; APPLICANT: Shelliga, Theodore A  
; TITLE OF INVENTION: Improved Methods for Processing Activated Protein C  
; FILE REFERENCE: APC process patent  
; CURRENT APPLICATION NUMBER: US/09/065,872  
; CURRENT FILING DATE: 1998-04-24  
; EARLIER APPLICATION NUMBER: 60/045,255  
; EARLIER FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 409  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-065-872-2  
Alignment Scores:  
Pred. No.: 2,52e-163 Length: 409  
Score: 2247.00 Matches: 405  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 96.15% Indels: 0  
DB: 4 Gaps: 0  
US-09-763-153-1 (1-1245) x US-09-065-872-2 (1-409)  
QY 31 AGCAGCGCTGGAGCGGAGTGCATAGAGAGATCTGTGACTTCGAGGAGGCAAGGAAT 90  
Db 1 SerSerLeuGluArgGluCysIleGluIleCysAspPheGluAlaLysGluIle 20  
QY 91 TTCAAAATGTGATGACACACTGGCTTCTGGTCCAAGCAGTGCAGCGTGACCACTGC 150  
Db 21 PheGlnAsnValAspAspThrLeuAlaPheTrpSerLysHisValAspGlyAspGlnCys 40  
QY 151 TTGGTCTTCCCTTGGAGCACCGCTGGCGCAGCTGTGCTGGCGGCGCGCTGCTGCTC 210  
Db 41 LeuValLeuProLeuGluHisProCysAlaSerLeuCysCysGlyHisGlyThrCysIle 60  
QY 211 GACGCGATCGCAGCTTCAGCTCGCAGCTCGCGCAGCGCTGGAGGCGCGCTTCTGCCAG 270  
Db 61 AspGlyIleGlySerPheSerCysAspCysArgSerGlyTyrPgluGlyArgPheCysGln 80  
QY 271 CGCAGGTGAGCTTCCTCAATTCCTGCTCGCTGGACAACCGCGGCTGCACGCAATTACTG 330  
Db 81 ArgGluValSerPheLeuAsnCysSerLeuAspAsnGlyGlyCysThrHisTyrCysLeu 100  
QY 331 GAGGAGTGGCTGGCGGCGCTGTAGCTGTGGCGCTGCATACAGTGGGCGGAGCACCTC 390  
Db 101 GluGluValGlyTyrArgArgCysSerCysAlaProGlyTyrLysLeuGlyAspLeu 120  
QY 391 CTCAGTGTCAACCCCGCAGTGAAGTTCCCTTGTGGAGGCGCTTGGAGGAGGATGAGAAG 450  
Db 121 LeuGlnCysHisProAlaValLysPheProCysGlyArgProTyrLysArgMetGluLys 140  
QY 451 AAGCGCAGTCACTGAACAGACACAGAGAACCAAGAACCAAGTAGATCGCGGCTC 510

```
Db 141 LysArgSerHisLeuLysArgAspThrGluAspGlnGluAspGlnValAspProArgLeu 160
QY 511 ATTGATGGGAAGATGACACAGCGGGGAGAGACACCCCTGGCAGGTGGTCTGCTGACACTCA 570
Db 161 IleAspGlyLysMetThrArgArgGlyAspSerProTrpGlnValValLeuLeuAspSer 180
QY 571 AAGAAGAAGCTGGCTGCGGGGAGTGGCTCATCCACCCCTCTGGTGGTGGTGGACAGCGCC 630
Db 181 LysLysLysLeuAlaCysGlyAlaValLeuIleHisProSerTrpValLeuThrAlaAla 200
QY 631 CACTGCATGGATGATCCCAAGAGCTCCTGTGACGGCTGGAGATGATGACTGCGGGCC 690
Db 201 HisCysMetAspGluSerLysLysLeuValArgLeuGlyGluTrpAspLeuArgArg 220
QY 691 TGGGAGAAGCTGGAGCTGGACCTGGACATCAGAGGAGTCTCGTCCACCCCAACTACAGC 750
Db 221 TrpGluLysTrpGluLeuAspLeuAspIleLysGluValPheValHisProAsnTrpSer 240
QY 751 AAGAGCACCACCGACAATGACATCGACTGCTGACCTGGCCCGCCCGCCACCCCTCTCG 810
Db 241 LysSerThrThrAspAsnAspIleAlaLeuLeuHisLeuAlaGlnProAlaThrLeuSer 260
QY 811 CAGACCATAGTCCCATCTGCTCCCGACAGCGCCCTTGACAGCGCGAGCTCAATCAG 870
Db 261 GlnThrIleValProIleCysLeuProAspSerGlyLeuAlaGluArgGluLeuAsnGln 280
QY 871 GCCGGCCAGGAGACCTCTGTCAGCGGCTGGGGCTACACAGCAGCGGAGAGAGGAGCC 930
Db 281 AlaGlyGlnGluThrLeuValThrGlyTrpGlyTrpHisSerArgGluLysGluAla 300
QY 931 AAGAGAAACCGCACCTTCTGCTCAACTTCATCAAGATTCCCGTGGTCCCGCCACAATGAG 990
Db 301 LysArgAsnArgThrPheValLeuAsnPheIleLysIleProValValProHisAsnGlu 320
QY 991 TGCAGCGAGCTCATGAGCAACATGTGTCTGAGAACATCTGTGTGGGGGATCCCTCGGG 1050
Db 321 CysSerGluValMetSerAsnMetValSerGluAsnMetLeuLysAlaGlyIleLeuGly 340
QY 1051 GACCGCAGGATGCTGCGAGGGGACAGTGGGGGCGCCATGGTGGCTCTCCACGGC 1110
Db 341 AspArgGlnAspAlaCysGluLysPheSerGlyGlyProMetValAlaSerPheHisGly 360
QY 1111 ACCTGGTCTGGTGGGCTGGTGGAGCTGGGGCTGGGGCTCTTCCAACTAC 1170
Db 361 ThrTrpPheLeuValGlyLeuValSerTrpGlyGluGlyCysGlyLeuLeuHisAsnTyr 380
QY 1171 GCGCTTTACACCAAGTACGCGCTTACCTCGACTGGATCCATGGCATCAGAGACAG 1230
Db 381 GlyValThrLysValSerArgTrpLeuAspTrpIleHisGlyHisIleArgAspLys 400
QY 1231 GAAGCCCCCAGAAG 1245
Db 401 GluAlaProGlnLys 405
```

## RESULT 17

US-09-667-570A-2

; Sequence 2, Application US/09667570A

; Patent No. 6436397

; GENERAL INFORMATION:

; APPLICANT: Baker, Jeffrey C

; APPLICANT: Carlson, Andrew D

; APPLICANT: Huang, Lihua

; APPLICANT: Sheliga, Theodore A

; TITLE OF INVENTION: Improved Methods for Processing Activated Protein C

; FILE REFERENCE: X-11796A

; CURRENT APPLICATION NUMBER: US/09/667,570A

; PRIOR FILING DATE: 2000-09-21

; PRIOR FILING DATE: 1997-04-28

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 409

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-667-570A-2
```

## Alignment Scores:

```
Pred. No.: 2,52e-163 Length: 409
Score: 2247.00 Matches: 405
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.15% Indels: 0
DB: 4 Gaps: 0
```

US-09-763-153-1 (1-1245) x US-09-667-570A-2 (1-409)

```
QY 31 AGCAGCTGGAGCGGAGTGCATAGAGAGATCTGTGACTTCGAGGAGGCCAAGAAATT 90
Db 1 SerSerLeuGluArgGluCysIleGluGluIleCysAspPheGluGluAlaLysGluIle 20
QY 91 TTCCAAAATGTGGATGACACACTGGCCCTTGTGTCCCAAGCAGCTCGACGGTGACCATGC 150
Db 21 PheGlnAsnValAspAspThrLeuAlaPheTrpSerLysHisValAspGlyAspGlnCys 40
QY 151 TTGCTCTTGGCCTTGGAGCACCCCTGCGCCAGCTGTGCTGCGGACAGCGCATGTGCATC 210
Db 41 LeuValLeuProLeuGluHisProCysAlaSerLeuCysCysGlyHisGlyThrCysIle 60
QY 211 GACGGCATCGGACCTTCAGCTGGGACTGCGGAGCGGCTGGGAGGGCCGCTTCTGGCAG 270
Db 61 AspGlyIleGlySerPheSerCysAspCysArgSerGlyTrpGluGlyArgPheCysGln 80
QY 271 CGCAGGTGAGCTTCCCAATTCCTGCTGCGACAAAGCGGCTGCAGCATTAATTACTGCCTA 330
Db 81 ArgGluValSerPheLeuAsnCysSerLeuAspAsnGlyGlyCysThrHisTyrCysLeu 100
QY 331 GAGAGGTGGGCTGGCGGCGCTGTAGCTGTGCGCTGCTACAGCTGGGGGAGCAGCCTC 390
Db 101 GluGluValGlyTrpArgArgCysSerCysAlaProGlyTrpLysLeuGlyAspAspLeu 120
QY 391 CTGCAGTGTCAACCCGAGTGAAGTCCCTTGTGGAGGCCCTGGGAAGCGGATGGAGAAG 450
Db 121 LeuGlnCysHisProAlaValLysPheProCysGlyArgProTrpLysArgMetGluLys 140
QY 451 AAGCGCAGTCACTGAAACGAGACACAGAACAGACCAAGACCAAGTAGATTCCCGGCTC 510
Db 141 LysArgSerHisLeuLysArgAspThrGluAspGlnGluAspGlnValAspProArgLeu 160
QY 511 ATTGATGGGAAGATGACCGGCGGAGACAGCCCTTGGCAGGTGTCTGTGCTGGACTCA 570
Db 161 IleAspGlyLysMetThrArgArgGlyAspSerProTrpGlnValLeuLeuAspSer 180
QY 571 AAGAAGAAGCTGGCGTGGGCGAGTGCATCCACCCCTCTGGTGTGACAGCGGCC 630
Db 181 LysLysLysLeuAlaCysGlyAlaValLeuIleHisProSerTrpValLeuThrAlaAla 200
QY 631 CACTGCATGGATGATCCAAAGAGCTCTTGTGAGGCTTGGAGAGTATGACCTCGCGCGC 690
Db 201 HisCysMetAspGluSerLysLysLeuValArgLeuGlyGluTrpAspLeuArgArg 220
QY 691 TGGGAGAAGTGGAGCTGGACCTGGACATCAAGAGGTCTTCCTCACCCCAACTACAGC 750
Db 221 TrpGluLysTrpGluLeuAspLeuAspIleLysGluValPheValHisProAsnTrpSer 240
QY 751 AAGAGCACCACCGACAATGACATCGCACTGCTGACCTGGCCCGCCCGCCACCCCTCTCG 810
Db 241 LysSerThrThrAspAsnAspIleAlaLeuLeuHisLeuAlaGlnProAlaThrLeuSer 260
QY 811 CAGACCATAGTGCCTCTGCTCCCGGAGCAGCGGCTTGCAGAGCGCGAGCTCAATCAG 870
Db 261 GlnThrIleValProIleCysLeuProAspSerGlyLeuAlaGluArgGluLeuAsnGln 280
QY 871 GCCGGCCAGGAGACCTCTGTGAGCGGCTGGGGCTACACAGCAGCGGAGAGGAGGCC 930
Db 281 AlaGlyGlnGluThrLeuValThrGlyTrpGlyTrpHisSerArgGluLysGluAla 300
```



```
QY 931 AAGAGAAACCGCACCTTCTGCTTCAACTTTCATCAAGATTCCCTGGTCCCGCACAAATGAG 990
|
|
|
Db 301 LysArgAsnArgThrPheValLeuAsnPheIleLysIleProValValProHisAsnGlu 320
|
|
|
QY 991 TGCAGCGAGGTCTAGCAACATGGTGTCTGAGAACATGCTGTGTGGGGCATCTCTCGG 1050
|
|
|
Db 321 CysSerGluValMetSerAsnMetValSerGluAsnMetLeuCysAlaGlyIleLeuGly 340
|
|
|
QY 1051 GACCGGACAGATGCTGGAGGCGACAGTGGGGGGCCCATGGTGCCTCTTCCACGGC 1110
|
|
|
Db 341 AspArgGlnAspAlaCysGluGlyAspSerGlyPrometValAlaSerPheHisGly 360
|
|
|
QY 1111 ACCTGGTCTCTGGTGGGCTGTGAGTGGGGTGGGGCTGTGGGCTCTTCAACATAC 1170
|
|
|
Db 361 ThrTrpPheLeuValGlyLeuValSerTrpGlyGluGlyCysGlyLeuLeuHisAsnTyr 380
|
|
|
QY 1171 GCGGTTTACAAAGTCAAGCGCTACCTGACTGAGTCCATGGGACATCAGAGACAAG 1230
|
|
|
Db 381 GlyValTyrThrLysValSerArgTyrLeuAspTrpIleHisGlyHisIleArgAspLys 400
|
|
|
1231 GAAGCCCCCAGAG 1245
|
|
|
401 GluAlaProGlnLys 405
|
|
|
RESULT 18
5270178-15
; Patent No. 5270178
; APPLICANT: GERLITZ, BRUCE E.; GRINNELL, BRIAN W.
; TITLE OF INVENTION: VECTORS AND COMPOUNDS FOR EXPRESSION OF
; ZYMOGEN FORMS OF HUMAN PROTEIN C
; NUMBER OF SEQUENCES: 21
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/484,133
; FILING DATE: 23-FEB-1990
; SEQ ID NO:15:
; LENGTH: 460
5270178-15

Alignment Scores:
Pred. No.: 3,37e-163 Length: 460
Score: 2245.50 Matches: 407
Percent Similarity: 98.80% Conservative: 3
Best Local Similarity: 98.07% Mismatches: 4
Query Match: 96.08% Indels: 1
Db: 6 Gaps: 1

US-09-763-153-1 (1-1245) x 5270178-15 (1-460)

QY 1 GCCAACTCTTCTGAGAGAGCTCGTCACAGACGCTTGAGCGGAGTGCATAGAGAG 60
|
|
|
Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGluGlu 62
|
|
|
61 ATCTGTGACTTCGAGGAGGCGCAAGAAATTTCCAAATGTTGGATGCACACTGGCCTTC 120
|
|
|
63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspThrLeuAlaPhe 82
|
|
|
QY 121 TGGTCCAAAGCAGCTCGACGGTGCACAGTGTGTGGTCTTGGCTTGGAGCACCGCTGGCGC 180
|
|
|
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102
|
|
|
QY 181 AGCTGTGCTGGGCGACGCGCATGTCATCGACGCAATCGGACGCTTCAGCTGGACTGC 240
|
|
|
Db 103 SerLeuLysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122
|
|
|
QY 241 CCGAGCGCTGGAGGCGGCTTCTGCCAGCGCGGAGTGCCTCAATTGCTCGCTG 300
|
|
|
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142
|
|
|
QY 301 GACAAAGCGCGCTGCACCAATCTACTGCTAGAGAGTGGGCTGGCGCGCTGTAGCTGT 360
|
|
|
Db 143 AspAsnGlyGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 162
|
|
|
```

```
QY 361 GCGCTTGGCTACAAAGCTGGGGAGCAGCCTCTGCAGTGTCAACCCCGCAGTGAAGTTCCT 420
|
|
|
Db 163 AlaProGlyTyrLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 182
|
|
|
QY 421 TGTGGAGGCGCTTGGAGCGGATGGAGAAGACGCGAGTCACTGAAACGACACAGAAA 480
|
|
|
Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202
|
|
|
QY 481 GACCAAGAAAGACCAAGTAGATCCGCGCTCATTTGATGGGAAGATGACACCGGGGAGAC 540
|
|
|
Db 203 AspGlnGluAspProValPheProArgLeu---AsnGlyLysMetThrArgArgGlyAsp 221
|
|
|
QY 541 AGCCCTTGGCAGGTGTCTGTGGACTCAAAGAGAAAGTGGCTGGCGGGGAGTGCCTC 600
|
|
|
Db 222 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 241
|
|
|
QY 601 ATCCACCCCTCTGGGTGTGTGACAGCGGCCCATCTGCATGATGATGAGTCAAGAGTCCCT 660
|
|
|
Db 242 IleHisProSerTrpValLeuThrAlaIleHisCysMetAspGluSerLysLeuIle 261
|
|
|
QY 661 GTCAGGCTTGGAGAGTATGACCTGCGCGCTGGGAGAGTGGGAGTGGACCTGGACATC 720
|
|
|
Db 262 ValArgIleGlyGluTyrAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 281
|
|
|
QY 721 AAGGAGGTCTTGTCCACCCCAACTACAGACAGAGCACCACGACATGACATCGACTG 780
|
|
|
Db 282 LysGluValPheValHisProAsnTyrSerLysSerThrAspAsnAspIleAlaLeu 301
|
|
|
QY 781 CTCGACCTGGCCCGACGCGCCCTCTCGACAGCATAGTCCCATCTGCCCGGAC 840
|
|
|
Db 302 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 321
|
|
|
QY 841 AGCGGCTTGCAGAGCGCGAGCTCAATCAGCGCGCGGAGGAGACCCCTCGTGCAGGCTGG 900
|
|
|
Db 322 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTyr 341
|
|
|
QY 901 GGCTACACAGCAGCGAGAGAGAGAGGAGGAGAAACCGACCTCTGCTCTCAACTTC 960
|
|
|
Db 342 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 361
|
|
|
QY 961 ATCAAGATTCCTGCTCCCGCACAAATGATGTCAGCGAGGTCTATGAGCAACATGGTGTCT 1020
|
|
|
Db 362 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 381
|
|
|
QY 1021 GAGAACATGCTGTGTGGGCGCATCTCGGCGGACCGCGAGGATGCTCGGAGGCGGACGT 1080
|
|
|
Db 382 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 401
|
|
|
QY 1081 GGGGGCGCATGTCGCTCTTCCAGCGCACCTGTTCTGTTGGGCTGCTGAGCTGG 1140
|
|
|
Db 402 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 421
|
|
|
QY 1141 GGTGAGGCTGTGGCTCTCTTCAAACTAGCGGCTTTTACACCAAGTACAGCGCTACCTC 1200
|
|
|
Db 422 GlyGluGlyCysGlyLeuLeuHisAsnValGlyValTyrThrHisValSerArgTyrLeu 441
|
|
|
QY 1201 GACTGGATCCATGGGCACATCAGAGACAAGAAAGAGCCCCCAGAG 1245
|
|
|
Db 442 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 456
|
|
|
RESULT 19
5270178-16
; Patent No. 5270178
; APPLICANT: GERLITZ, BRUCE E.; GRINNELL, BRIAN W.
; TITLE OF INVENTION: VECTORS AND COMPOUNDS FOR EXPRESSION OF
; ZYMOGEN FORMS OF HUMAN PROTEIN C
; NUMBER OF SEQUENCES: 21
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/484,133
; FILING DATE: 23-FEB-1990
; SEQ ID NO:16:
; LENGTH: 460
5270178-16
```



Alignment Scores:

Pred. No.: 3.81e-161 Length: 460  
Score: 2218.50 Matches: 403  
Percent Similarity: 98.07% Conservative: 7  
Best Local Similarity: 97.11% Mismatches: 4  
Query Match: 94.93% Indels: 1  
DB: 6 Gaps: 1

US-09-763-153-1 (1-1245) x 5270178-16 (1-460)

QY 1 GCCAAGTCTCTCTGAGGAGTCCGTCACAGCAGCCTGGAGCGGAGTGCATAGGAG 60  
43 AlaAsnSerPheLeuGluLeuArgHisSerSerLeuGluArgGluCysIleGlu 62  
61 ATCTGTACTTCGAGAGGCCAAGGAAATTTCCAAATGTGGATGACACACTGGCCCTC 120  
63 IleCysAspPheGluAlaLysGluIlePheGlnAsnValAspThrLeuAlaPhe 82  
QY 121 TGGTCCAAGCAGTCGACGGTGACAGTCTTGGTCTTGGCCCTTGGAGCACCCTGGCC 180  
83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
QY 181 AGCCTGTCTGCGGGCAGCGCAGTGCATCGACGCGCATCGGAGCTTCAGCTGCCACTGC 240  
103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
QY 241 CGCAGCGCTGGAGGGCGCTCTCTGCCAGCGAGGTAGCTTCTCAATTGCTCGGTG 300  
123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
QY 301 GACAAGCGGCTGACGACGATCTACTGCTAGAGGAGTGGGCTGGCGGCTGTAGCTGT 360  
143 AspAsnGlyCysThrHisTyrCysLeuGluValGlyTrpArgArgCysSerCys 162  
QY 361 GCCTCTGCTACAAGTGGGAGCGACCTCTCTGCACTGTCACCGCGAGTGAAGTTCCT 420  
163 AlaProGlyTyrLysLeuGlyAspLeuLeuGlnCysHisProAlaValLysPhePro 182  
QY 421 TGTGGAGGCGCTGGAGCGGATGGAGAGAGCGGAGTCACTGAAAGCAGACAGAA 480  
183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
QY 481 GACCAAGAGACCAACTAGTACCGCGCTCATTGATGGAAGTACACAGCGGGGAGAC 540  
203 LeuHisLysLeuGlnThrTyrProArgLeu---AsnGlyLysMetThrArgGlyAsp 221  
541 AGCCCTGGCAGGTGCTCTGCTGGACTCAAGAAGAGCTGGCTGGCGGCGAGTGTCT 600  
222 SerProTrpGlnValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 241  
QY 601 ATCCACCTCTCTGGGTGTGTGACAGCGGCCACTGCTGATGATGAGTCCAAAGCTCTCT 660  
242 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuIle 261  
QY 661 GTGAGCTTGGAGATGACCTGCGCGCTGGAGAGTGGAGTGGAGTGGACCTGGACATC 720  
262 ValArgIleGlyGlyTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 281  
QY 721 AAGGAGGTCTCTGCCACCCCACTACAGCAGAGCACCACCGCAATGACATCGCAGCTG 780  
282 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 301  
QY 781 CTCACCTGGCCCGCCGCCCTCTCGCAGACATAGTCCCATCTGCTCCCGGAC 840  
302 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 321  
QY 841 AGCGGCTTTCAGAGCGCGAGTCAATCAGCGCGGCCAGGAGACCTCTGTCGCGGCTGG 900  
322 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 341  
QY 901 GGTACACAGCAGCGAGAGAGGAGGCGCAAGAGAAACCGGACCTCTGCTCTCAACTTC 960

Db 342 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 361  
QY 961 ATCAAGATTCCTGGTCCCGCACAAATGAGTGCAGCGAGTGCATGACACATGGTGTCT 1020  
Db 362 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 381  
QY 1021 GAGAACATGCTGTGGGGCATCTCTCGGGAGCGGCGAGGATGCTCGAGGGGCGACAGT 1080  
Db 382 GluAsnMetLeuLysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 401  
QY 1081 GGGGGGCCATGCTGCTCCACGCGACCTCTCTCGGGAGCGGCGAGTGCATGAGCTGG 1140  
Db 402 GlyGlyPrometValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 421  
QY 1141 GGTGAGGCTGTGGGCTCTCTCACAACACTAGCGGCTTTACACCAAGTACGCCGCTACTC 1200  
Db 422 GlyGluGlyCysGlyLeuLeuHisAsnValGlyValThrHisValSerArgTyrLeu 441  
QY 1201 GACTGGATCCATGGGCACATCAGACAGCAAGGAGCGGCGGCGAGG 1245  
Db 442 AspIleHisGlyHisIleArgAspLysGluAlaProGlnLys 456

RESULT 20

US-07-720-189-1  
; Sequence 1, Application US/07720189  
; Patent No. 5279956

; GENERAL INFORMATION:  
; APPLICANT: Griffin, John H.  
; APPLICANT: Mesters, Rolf M.  
; TITLE OF INVENTION: APC POLYPEPTIDES AND ANTI-PEPTIDE  
; TITLE OF INVENTION: ANTIBODIES, DIAGNOSTIC METHODS AND SYSTEMS FOR INHIBITING  
; TITLE OF INVENTION: APC, AND THERAPEUTIC METHODS  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute, Office of Patent  
; ADDRESSEE: Counsel  
; STREET: 3366 No. 5279956th Torrey Pines Court, Suite 240  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07720189  
; FILING DATE: 19910724  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bingham, Douglas  
; REGISTRATION NUMBER: 32,457  
; REFERENCE/DOCKET NUMBER: SCRO390P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-554-2937  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 262 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Region  
; LOCATION: 1..262  
; OTHER INFORMATION:  
; OTHER INFORMATION: the PC heavy chain, the amino acid residue positions of  
; OTHER INFORMATION: which begin at position 158 and end at 419."

US-07-720-189-1

## Alignment Scores:

Pred. No.:	2.06e-98	Length:	262
Score:	1393.00	Matches:	258
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	59.61%	Indels:	0
DB:	1	Gaps:	0

US-09-763-153-1 (1-1245) x US-07-720-189-1 (1-262)

QY	472	GACACAGAACCAAGACCAAGACAGATGATCGCGGCTATTGATGGGAAGATGACCAAG	531
DB	1	AspThrGluAspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArg	20
QY	532	CGGGAGACAGCCCTCGGAGGTGCTGCTGGACTCAAAAGAAAGCTGGCCTGCGGG	591
DB	21	ArgGlyAspSerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGly	40
QY	592	GCAGTCTCATCCACCCCTCTCGGTGTGTGACAGGGCCCACTGCATGGATGAGTCCAAG	651
DB	41	AlaValLeuIleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLys	60
QY	652	AAGCTCTCTGTGAGGTGGAGAGTATGACCTGGCGCTGGAGAGAGTGGGAGCTGGAC	711
DB	61	LysLeuLeuValArgLeuGlyGluTrpAspLeuArgArgTrpGluLysTrpGluLeuAsp	80
QY	712	CTGGACATCAAGAGGTCTCTCCACCCCACTACAGAACAGACACCCGACCAATGAC	771
DB	81	LeuAspIleLysGluValPheValHisProAsnTyrSerLysSerThrAspAsnAsp	100
QY	772	ATCGCACTGCTCCACCTGGCCGCCGCCACCTCTCGGAGACCATAGTCCCATCTGC	831
DB	101	IleAlaLeuLeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCys	120
QY	832	CTCCGGACAGCGGCTTTCAGAGCGGAGCTCAATCAGCGCGCCAGGAGACCCCTCGTG	891
DB	121	LeuProAspSerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuVal	140
QY	892	ACGGCTGGGGCTACACAGAGCGGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG	951
DB	141	ThrGlyTrpGlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheVal	160
QY	952	CTCACTTCATCAGATTCCTCGGTGCTCCCGCACATGAGTGCAGGAGGTCATGAGCAAC	1011
DB	161	LeuAsnPheIleLysIleProValProHisAsnGluCysSerGluValMetSerAsn	180
QY	1012	ATGTTGTGTGAGACATGCTGTGCGGGCATCTCGGGGACCGGAGGATGCTGCGAG	1071
DB	181	MetValSerGluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGlu	200
QY	1072	GGCGACGTGGGGCCCATGCTGCTCTCCCTCCACGCGACCTGGTTCCTGGTGGGCTG	1131
DB	201	GlyAspSerGlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeu	220
QY	1132	GTGAGCTGGGTGAGGGCTGGGCTCTCTCACAACTACGCGCTTTACACCAAGTCACG	1191
DB	221	ValSerTrpGlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSer	240
QY	1192	CGCTACCTCGACTGGATTCATGGGCACATCAGACACAAAGGAGGAGGAGGAGGAGGAG	1245
DB	241	ArgTyrLeuAspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys	258

Search completed: March 17, 2003, 17:38:42  
Job time : 56 secs

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: March 17, 2003, 17:35:25 ; Search time 47 Seconds  
(without alignments)  
2441.899 Million cell updates/sec

Title: US-09-763-153-1

Perfect score: 2337

Sequence: 1 gccaaactctctctggagga.....acaaggaagccccccagaag 1245

oring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 398832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 60 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool/US09763153/runat\_11032003\_084249\_2120/app\_query.fasta\_1.1415  
-DB=PublishedApplications\_AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62  
-TRANS=human40.cdi -LIST=60 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=20 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09763153.acgn.1.1.12 @runat\_11032003\_084249\_2120  
-NCPU=6 -ICPU=3 -NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pcp.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pcp.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pcp.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pcp.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pcp.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pcp.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pcp.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pcp.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pcp.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2298	98.3	419	9	US-10-182-263-1
2	2298	98.3	419	9	US-09-978-917A-4
3	2298	98.3	461	9	US-10-182-263-2
4	2298	98.3	461	9	US-09-978-917A-2

#### ALIGNMENTS

US-10-182-263-1

; Sequence 1, Application US/10182263  
; Publication No. US20030022354A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerlitz, Bruce E  
; APPLICANT: Jones, Bryan E  
; APPLICANT: Grinnell, Brian W  
; TITLE OF INVENTION: PROTEIN C DERIVATIVES  
; FILE REFERENCE: X-13611  
; CURRENT APPLICATION NUMBER: US/10/182,263  
; CURRENT FILING DATE: 2002-07-22  
; PRIOR APPLICATION NUMBER: 60/181948

5	2270	97.1	419	9	US-10-182-263-5	Sequence 5, Appli
6	2264	96.9	419	9	US-10-182-263-3	Sequence 3, Appli
7	2262	96.8	419	9	US-10-182-263-6	Sequence 6, Appli
8	2260	96.7	419	9	US-10-182-263-4	Sequence 4, Appli
9	736	31.5	461	9	US-10-132-829-5	Sequence 5, Appli
10	736	31.5	461	10	US-09-884-901-3	Sequence 3, Appli
11	735	31.5	415	10	US-09-118-748-2	Sequence 1, Appli
12	726	31.1	406	9	US-10-109-498-1	Sequence 1, Appli
13	470	20.1	802	9	US-09-978-295A-169	Sequence 169, App
14	470	20.1	802	9	US-09-978-697-169	Sequence 169, App
15	470	20.1	802	9	US-09-978-192A-169	Sequence 169, App
16	470	20.1	802	9	US-09-999-832A-169	Sequence 169, App
17	470	20.1	802	9	US-09-978-189-169	Sequence 169, App
18	470	20.1	802	9	US-09-978-608A-169	Sequence 169, App
19	470	20.1	802	10	US-09-888-615-113	Sequence 113, App
20	428.5	18.3	850	10	US-09-888-615-108	Sequence 108, App
21	426.5	18.2	249	9	US-09-961-721-5	Sequence 5, Appli
22	426.5	18.2	251	9	US-09-961-721-4	Sequence 4, Appli
23	424.5	18.2	638	9	US-09-808-602-102	Sequence 102, App
24	420	18.0	679	10	US-09-874-238-6	Sequence 6, Appli
25	419.5	18.0	812	9	US-09-335-325-1	Sequence 1, Appli
26	419.5	18.0	812	9	US-10-131-241-1	Sequence 1, Appli
27	419.5	18.0	812	10	US-09-788-142-1	Sequence 1, Appli
28	419.5	18.0	812	10	US-09-761-120-1	Sequence 1, Appli
29	419.5	18.0	812	10	US-09-873-676-81	Sequence 81, Appli
30	419.5	18.0	625	9	US-09-808-602-104	Sequence 104, App
31	419	17.9	643	9	US-09-808-602-103	Sequence 103, App
32	417.5	17.9	855	10	US-09-900-751-2	Sequence 2, Appli
33	417.5	17.9	560	10	US-09-912-559-3	Sequence 3, Appli
34	410	17.5	791	9	US-09-967-386-1	Sequence 1, Appli
35	410	17.5	810	10	US-09-946-893-2	Sequence 2, Appli
36	410	17.5	822	9	US-09-147-947-6	Sequence 6, Appli
37	408.5	17.4	227	9	US-10-045-367A-4	Sequence 4, Appli
38	406.5	17.4	393	9	US-10-012-896-934	Sequence 934, App
39	405.5	17.4	393	9	US-09-895-793-934	Sequence 934, App
40	405.5	17.4	393	9	US-09-895-814-934	Sequence 934, App
41	405.5	17.4	393	10	US-09-759-143-934	Sequence 934, App
42	405.5	17.4	393	10	US-09-780-669-934	Sequence 934, App
43	405.5	17.4	393	10	US-09-822-827-934	Sequence 934, App
44	405.5	17.4	492	9	US-10-012-896-932	Sequence 932, App
45	405.5	17.4	492	9	US-09-895-793-932	Sequence 932, App
46	405.5	17.4	492	9	US-09-895-814-932	Sequence 932, App
47	405.5	17.4	492	10	US-09-759-143-932	Sequence 932, App
48	405.5	17.4	492	10	US-09-780-669-932	Sequence 932, App
49	405.5	17.4	492	10	US-09-822-827-932	Sequence 932, App
50	405.5	17.4	384	9	US-09-981-353-23	Sequence 23, Appli
51	402.5	17.2	492	9	US-10-012-896-895	Sequence 895, App
52	402.5	17.2	492	9	US-09-895-793-895	Sequence 895, App
53	402.5	17.2	492	9	US-09-895-814-895	Sequence 895, App
54	402.5	17.2	492	10	US-09-759-143-895	Sequence 895, App
55	402.5	17.2	492	10	US-09-780-669-895	Sequence 895, App
56	402.5	17.2	492	10	US-09-822-827-895	Sequence 895, App
57	402.5	17.2	560	10	US-09-912-559-4	Sequence 4, Appli
58	402.5	17.2	560	10	US-09-912-559-4	Sequence 4, Appli
59	402	17.2	855	9	US-10-099-700A-2	Sequence 2, Appli
60	401.5	17.2				

```
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-1

Alignment Scores:
Pred. No.: 1,48e-127 Length: 419
Score: 2298.00 Matches: 415
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.33% Indels: 0
DB: Gaps: 0

US-09-763-153-1 (1-1245) x US-10-182-263-1 (1-419)
1 GCCAACTCCTCTGTGAGGAGCTCCGTACAGCAGCTGGAGCGGAGTGCATAGAGGAG 60
1 AlaAsnSerPheLeuGluGluLeuArgHisSerLeuGluArgGluCysIleGlu 20
61 ATCTGTGACTTCCGAGGAGCCCAAGGAAATTTCCAAATGTGGATCACACACTGGCCTTC 120
21 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 40
121 TGTCTCAAGCAGCTGACGGGTGACAGTGTGCTGTGCTTCCCTTGGAGCACCCTGGCGCC 180
41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60
181 AGCTGTGCTGCGGCGCAGCGCATGCTGCATCGAGCGCATCGGAGCTTCAGCTGCAGCTGC 240
61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80
241 CGCAGCGGCTGGAGGCGCGCTTCTGCCAGCGGAGGTGAGTTCCTCAATGTCTGCTG 300
81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100
301 GACACGCGGCTGACGAGTACTGCTAGAGGAGTGGCTGGCGGCGCTGTAGCTCT 360
101 AspAsnGlyGlyCysThrHisThrCysLeuGluGluValGlyTrpArgCysSerCys 120
361 GCGCTGCTACAGCTGGGGGAGCAGCTCTCTGAGTGTACCCCGCAGTGAAGTTCCT 420
121 AlaProGlyTrpLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 140
421 TGTGGGAGCCCTGGAGCGGATGGAGAGAACGCGAGTCACCTGAAACGAGACACAGAA 480
141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160
481 GACCAAGACACCAAGTAGATCCGCGCTCATTTGATGGAGAGATGACAGCGCGGAGAC 540
161 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 180
541 AGCCCTCGGAGGTGCTGCTGCTGAGCTCAAGAAGAGCTGCGCTGCGGGGAGTGCCTC 600
181 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 200
601 ATCCACCCCTCTGGTGTGACAGCGGCGCCACTGATGATGATGATGATGATGATGATGAT 660
201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 220
661 GTCAGGCTTGGAGAGTATGACCTGCGCGCTGGGAGAGTGGAGAGTGGAGAGTGGAGAGT 720
221 ValArgLeuGlyGluTrpAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 240
721 AAGGAGGTCTTCTGCTCCACCCCACTACAGCAAGAGCACCACCGACAAATGATCGACTG 780
241 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 260

; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-1

Alignment Scores:
Pred. No.: 1,48e-127 Length: 419
Score: 2298.00 Matches: 415
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.33% Indels: 0
DB: Gaps: 0

US-09-763-153-1 (1-1245) x US-09-763-153-1 (1-419)
1 GCCAACTCCTCTGTGAGGAGCTCCGTACAGCAGCTGGAGCGGAGTGCATAGAGGAG 60
1 AlaAsnSerPheLeuGluGluLeuArgHisSerLeuGluArgGluCysIleGlu 20
61 ATCTGTGACTTCCGAGGAGCCCAAGGAAATTTCCAAATGTGGATCACACACTGGCCTTC 120
21 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 40
121 TGTCTCAAGCAGCTGACGGGTGACAGTGTGCTGTGCTTCCCTTGGAGCACCCTGGCGCC 180
41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60
181 AGCTGTGCTGCGGCGCAGCGCATGCTGCATCGAGCGCATCGGAGCTTCAGCTGCAGCTGC 240
61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80
241 CGCAGCGGCTGGAGGCGCGCTTCTGCCAGCGGAGGTGAGTTCCTCAATGTCTGCTG 300
81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100
301 GACACGCGGCTGACGAGTACTGCTAGAGGAGTGGCTGGCGGCGCTGTAGCTCT 360
101 AspAsnGlyGlyCysThrHisThrCysLeuGluGluValGlyTrpArgCysSerCys 120
361 GCGCTGCTACAGCTGGGGGAGCAGCTCTCTGAGTGTACCCCGCAGTGAAGTTCCT 420
121 AlaProGlyTrpLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 140
421 TGTGGGAGCCCTGGAGCGGATGGAGAGAACGCGAGTCACCTGAAACGAGACACAGAA 480
141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160
481 GACCAAGACACCAAGTAGATCCGCGCTCATTTGATGGAGAGATGACAGCGCGGAGAC 540
161 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 180
541 AGCCCTCGGAGGTGCTGCTGCTGAGCTCAAGAAGAGCTGCGCTGCGGGGAGTGCCTC 600
181 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 200
601 ATCCACCCCTCTGGTGTGACAGCGGCGCCACTGATGATGATGATGATGATGATGATGAT 660
201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 220
661 GTCAGGCTTGGAGAGTATGACCTGCGCGCTGGGAGAGTGGAGAGTGGAGAGTGGAGAGT 720
221 ValArgLeuGlyGluTrpAspLeuArgArgTrpGluLysTrpGluLeuAspLeuAspIle 240
721 AAGGAGGTCTTCTGCTCCACCCCACTACAGCAAGAGCACCACCGACAAATGATCGACTG 780
241 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 260
```

```
Db 61 SerLeuCysCysGlyHisGlyThrCysLeuAspGlyLeuGlySerPheSerCysAspCys 80
QY 241 CGCAGCGCTGGAGGCGCGCTTCTGCGAGCGAGTGAGTCTCTCAATTCCTGCTGCTG 300
Db 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100
QY 301 GACACGCGCTGCTGACGCAATCTACTGCTAGAGGAGTGGCTGGCGGCGCTGACTGT 360
Db 101 AspAsnGlyGlyCysThrHisTyrcysLeuGluGluValGlyTrpArgGlySerCys 120
QY 361 GCGCTGCTACAAGCTGGGAGCGACTCTCTGCTAGTGTACCCCGCAGTGAAGTTCCT 420
Db 121 AlaProGlyTrpLysLeuGlyAspLeuGluCysHisProAlaValLysPhePro 140
QY 421 TGTGGAGCGCTGGAAGCGGATGGAGAACGCGGAGTCACTGAGTGGGAGTGGAGTGG 480
Db 141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160
QY 481 GACCAAGAACCAAGTAGATCGCGCTCATTTGATGGGAAGATGACACGCGGGAGAC 540
Db 161 AspGlnGluAspGlnValAspProArgLeuLeuAspGlyLysMetThrArgGlyAsp 180
QY 541 AGCCCTGCGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 181 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 200
QY 601 ATCCACCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLeuLeu 220
QY 661 GTGAGCTGGAGATGATGACCTGGCGCTGGGAGTGGGAGTGGGAGTGGGAGTGG 720
Db 221 ValArgLeuGlyLysTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 240
QY 721 AAGAGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 241 LysGluValPheValHisProAsnTrpSerLysSerThrThrAspAsnAspIleAla 260
QY 781 CTGACCTGGCGAGCGCGCCCTCTCGCAGACCATAGTGGCCATCTGCTGCTGCTGCT 840
Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProLysLeuProAsp 280
QY 841 AGCGGCTTGCAGAGCGGCTCAATCAGCGCGCGCGAGGAGCCCTGTCGAGCGGCTG 900
Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGly 920
QY 901 GGCTACACAGAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
Db 301 GlyTrpHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsn 320
QY 961 ATCAAGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 321 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetVal 340
QY 1021 GAGAACATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 341 GluAsnMetLeuCysAlaGlyLysLeuGlyAspArgGlnAspAlaCysGluGlyAsp 360
QY 1081 GGGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSer 380
QY 1141 GGTGAGGCTGTGGCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValThrLysValSerArgTrpLeu 400
QY 1201 CACTGGATCCATGGGACATCAGACAAAGGAAAGCCCGCCAGAG 1245
Db 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415
```

RESULT 3

US-10-182-263-2

```
; Sequence 2, Application US/10182263
; Publication No. US20030022354A1
; GENERAL INFORMATION:
; APPLICANT: Gerlitz, Bruce E
; APPLICANT: Jones, Bryan E
; APPLICANT: Grinnell, Brian W
; TITLE OF INVENTION: PROTEIN C DERIVATIVES
; FILE REFERENCE: X-13611
; CURRENT APPLICATION NUMBER: US/10/182,263
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 60/181948
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 60/189199
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-263-2
```

## Alignment Scores:

```
Pred. No.: 1.49e-127 Length: 461
Score: 2298.00 Matches: 415
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.33% Indels: 0
DB: 9 Gaps: 0
```

US-09-763-153-1 (1-1245) x US-10-182-263-2 (1-461)

```
QY 1 GCCAATCTCTCTGAGGAGCTCCGTCACAGCAGCTGGAGCGGAGTGCATAGAGAG 60
Db 43 AlaAsnSerPheLeuGluGluLeuArgHisSerSerLeuGluArgGluCysIleGlu 62
QY 61 ATCTGTGACTTCGAGAGGAGCCAAAGAAATTTCCAAAATGTGGATGACACACTGCCTTC 120
Db 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82
QY 121 TGGTCAAGCAGCTCGACGGTGGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102
QY 181 AGCTGTGCTGGGCGACGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122
QY 241 CGCAGCGCTGGGAGCGCGCTTCTGCGAGCGGAGTGGGCTGCTGCTGCTGCTGCTGCT 300
Db 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142
QY 301 GACACGCGCTGCTGACGCAATCTACTGCTAGAGGAGTGGGCTGCTGCTGCTGCTGCTGCT 360
Db 143 AspAsnGlyGlyCysThrHisTyrcysLeuGluGluValGlyTrpArgGlySerCys 162
QY 361 GCGCTGCTACAAGCTGGGAGCGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 163 AlaProGlyTrpLysLeuGlyAspLeuGluCysHisProAlaValLysPhePro 182
QY 421 TGTGGAGCGCTGGAAGCGGATGGAGAACGCGGAGTCACTGAGTGGGAGTGGAGTGG 480
Db 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202
QY 481 GACCAAGAACCAAGTAGATCGCGCTCATTTGATGGGAAGATGACACGCGGGAGAC 540
Db 203 AspGlnGluAspGlnValAspProArgLeuLeuAspGlyLysMetThrArgGlyAsp 222
QY 541 AGCCCTGCGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 223 SerProTrpGlnValValLeuLeuAspSerLysLysLeuAlaCysGlyAlaValLeu 242
QY 601 ATCCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
```

DB: 9 Gaps: 0  
US-09-763-153-1 (1-1245) x US-09-978-917A-2 (1-461)  
QY 1 GCCAACTCTCTCTGAGGAGTCTCTACAGACGCTGGAGCGGAGTGCATAGAGAG 60  
DB 43 AlaAsnSerPheLeuGluGluArgHisSerSerLeuGluArgGluCysIleGluGlu 62  
QY 61 ATCTGTGACTTCGAGGAGGCAAGGAAATTTCCAAATATGTGGATGACACACTGCCCTTC 120  
DB 63 IleCysAspPheGluGluAlaLysGluIlePheGlnAsnValAspAspThrLeuAlaPhe 82  
QY 121 TGGTCCAAAGCAGCTCGACGCTGACAGTGTGGTCTTGGCTTGGACACCGCTGCCGC 180  
DB 83 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 102  
QY 181 AGCTGTGCTGCGGCGACGCGACGCTGATCGACGCGATCGGAGCTTCAGCTGCAGTGC 240  
DB 103 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 122  
QY 241 CGCAGCGCTGGGAGGCGCGCTTCTGCCAGCGGAGGTGAGTTCCTCAATTCCTGCTG 300  
DB 123 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 142  
QY 301 GACACGCGCTGCGACGCTGATCTGCTAGAGGAGTGGCTGGCGGCGCTGTAGCTGT 360  
DB 143 AspAsnGlyGlyCysThrHisThrCysIleAspGlyIleGlySerPheSerCysAspCys 162  
QY 361 GCGCTGCTGCTACAGCTGGGAGCGACCTCTCTGCAAGTGTGACCCCGAGTGAAGTTCCT 420  
DB 163 AlaProGlyTrpLysLeuGlyAspLeuLeuGlnCysHisProAlaValLysPhePro 182  
QY 421 TGTGGAGGCGCTGGAACGCGATGGAGAGAGCGAGTACCTGAAACGAGACACAGAA 480  
DB 183 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 202  
QY 481 GACCAAGAGACCAAGTAGTCCGGGCTCATTTGATGGAAAGATGACACGCGGGAGAC 540  
DB 203 AspGlnGluAspGlnValAspProArgLeuIleAspGlyLysMetThrArgArgGlyAsp 222  
QY 541 AGCCCTGCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
DB 223 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyValaLeu 242  
QY 601 ATCCACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
DB 243 IleHisProSerTrpValLeuThrAlaHisCysMetAspGluSerLysLysLeuLeu 262  
QY 661 GTCAGGCTTGGAGAGTATGACCTGCGGCGCTGGGAGAGTGGGAGCTGGACCTGGACATC 720  
DB 263 ValArgLeuGlyGluTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
QY 721 AAGGAGTCTTCTGCTCCACCCCACTACAGCAAGAGCACCACCAATGACATCGACTG 780  
DB 283 LysGluValPheValHisProAsnTrpSerLysSerThrThrAspAsnAspIleAlaLeu 302  
QY 781 CTGCACCTGCGCGCGCGCGCGCTCTCTCCAGACCATAGTCCCATCTGCTCCCGGAC 840  
DB 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
QY 841 AGCGGCTTTCAGAGCGCGAGCTCATATCAGGCGCGCGCGAGAGACCTCTGACGGGCTGG 900  
DB 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
QY 901 GGCTTACCACAGCAGCGAGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
DB 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
QY 961 ATCAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
DB 363 IleLysIleProValProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 382  
QY 1021 GAGAACATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080

DB 243 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 262  
QY 661 GTCAGGCTTGGAGAGTATGACTGCGGCGCTGGGAGAACTGGAGCTGGACCTGGACATC 720  
DB 263 ValArgLeuGlyGluTrpAspLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 282  
QY 721 AAGGAGTCTTCTGCTCCACCCCACTACAGCAAGAGCACCACCAATGACATCGACTG 780  
DB 283 LysGluValPheValHisProAsnTrpSerLysSerThrThrAspAsnAspIleAlaLeu 302  
QY 781 CTGCACCTGCGCGCGCGCGCGCTCTCCAGACCATAGTGCCTGCTGCCCTGCCGAC 840  
DB 303 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 322  
QY 841 AGCGGCTTTCAGAGCGCGAGCTCAATCAGCGCGCGCGAGAGACCTCTGAGCGGCTGG 900  
DB 323 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 342  
QY 901 GGCTACCACAGCAGCGGAGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
DB 343 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 362  
QY 961 ATCAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
DB 363 IleLysIleProValProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 382  
QY 1021 GAGAACATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
DB 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
QY 1081 GGGGGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
DB 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
QY 1141 GGTGAGGCTGTGGGCTCTTCCACAACTACGCGCTTTACACAAAGTACGCGCTACCTC 1200  
DB 423 GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValTrpThrLysValSerArgTrpLeu 442  
QY 1201 GACTGATCATGCGGACATCAGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1245  
DB 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

RESULT 4  
US-09-978-917A-2  
; Sequence 2, Application US/09978917A  
; Publication No. US20030027299A1  
; GENERAL INFORMATION:  
; APPLICANT: Maxygen Aps; Maxygen Holdings  
; TITLE OF INVENTION: Protein C or activated protein C-like molecules  
; FILE REFERENCE: 0219us310 - protein C  
; CURRENT APPLICATION NUMBER: US/09/978,917A  
; CURRENT FILING DATE: 2001-10-17  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 461  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)...(42)  
; FEATURE:  
; NAME/KEY: CHAIN  
; LOCATION: (43)...(461)  
US-09-978-917A-2  
  
Alignment Scores:  
Pred. No.: 1,49e-127 Length: 461  
Score: 2298.00 Matches: 415  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.33% Indels: 0

Db 383 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 402  
QY 1081 GGGGGCCCATGGTCCCTCCCTCCACGGCAGCTGTTCCCTGGTGGGCTGTGAGCTGG 1140  
Db 403 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 422  
QY 1141 GGTGAGGCTGTGGGCTCTTCCAACTACGGCGTTTACACCAAGTCAGCCGCTAGCTC 1200  
Db 423 GlyGluGlyCysGlyLeuLeuHisAsnTyrglyValTyThrLysValSerArgTyLeu 442  
QY 1201 GACTGATCCATGGGCACATCAGACAAAGAAAGCCGCCCAAGAAG 1245  
Db 443 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 457

## MULT 5

10-182-263-5  
; Sequence 5, Application US/10182263  
; Publication No. US20030022354A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerlitz, Bruce E  
; APPLICANT: Jones, Bryan E  
; APPLICANT: Grinnell, Brian W  
; TITLE OF INVENTION: PROTEIN C DERIVATIVES  
; FILE REFERENCE: X-13611  
; CURRENT APPLICATION NUMBER: US/10/182,263  
; PRIORITY FILING DATE: 2002-07-22  
; PRIOR FILING DATE: 2002-02-11  
; PRIOR APPLICATION NUMBER: 60/189199  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 419  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-182-263-5

## Alignment Scores:

Pred. No.: 6,48e-126 Length: 419  
Score: 2270.00 Matches: 410  
Percent Similarity: 99.28% Conservative: 2  
Best Local Similarity: 98.80% Mismatches: 3  
Query Match: 97.13% Indels: 0  
DB: 9 Gaps: 0

us-09-763-153-1 (1-1245) x US-10-182-263-5 (1-419)

QY 1 GCCAACTCCCTTCCTGGAGAGCTCCGTCACAGCCCTGGAGCGGAGTGATAGAGAG 60  
Db 1 AlaAsnSerPheLeuGluGluLeuArgHisGlySerLeuGluArgGluCysIleGlu 20  
QY 61 ATCTGTGACTTCGAGAGGCCAAGGAAATTTTCCAAATGTGGATGACACACTGGCCCTC 120  
Db 21 IleCysAspPheGluGluAlaLysGluIlePheGluAspValAspThrLeuAlaPhe 40  
QY 121 TGTCTCAAGCAGCTGACGAGTCCAGCTGCTGGTCTTGGCTTGGAGCAGCCGCTGGCC 180  
Db 41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProGlnHisProCysAla 60  
QY 181 AGCTGTGTGGGGCAGCGCAGCTGATCGACGGCATCGGAGCTTCAGCTGCGACTGC 240  
Db 61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
QY 241 CGCAGCGGCTGGAGCGGCTTCGCGAGCGGAGGTGAGCTTCCTCAATGCTCGCTG 300  
Db 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
QY 301 GACAAGCGGCTGACGACCATCTACTGCTAGAGAGGTGGCTGGCGGCTGTAGCTGT 360  
Db 101 AspAsnGlyGlyCysThrHisTyrcysLeuGluGluValGlyTrpArgCysSerCys 120

QY 361 GGCCTCGCTACAAGCTGGGGACGACCTCTCTGCACTGTCAACCCGCGAGTGAAGTCCCT 420  
Db 121 AlaProGlyTyrlsLeuGlyAspAspLeuGlnCysHisProAlaValLysPhePro 140  
QY 421 TGTGGAGGCCCTTGAAGCGGATGAGAGAGCGCAGTCACTACCTGAAACGAGACACAGAA 480  
Db 141 CysGlyArgProTrpLysArgMetGluLysLysArgSerHisLeuLysArgThrGlu 160  
QY 481 GACCAAGAGACCAAGTAGATCCCGGCTCATTTGATGGGAAGATGACACGCGGGAGAC 540  
Db 161 AspGlnGluAspGlnValPheProArgLeuIleLysGlyLysMetThrArgArgGlyAsp 180  
QY 541 AGCCCTCGCAGGTGCTCTGCTGCTCAAGAGAGCTGGCTCGGGGAGCTGCTC 600  
Db 181 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 200  
QY 601 ATCCACCCCTCTCTGGTGTCTGACAGCGGCCCTGCTGATGATGATGATGATGATGATG 660  
Db 201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeuLeu 220  
QY 661 CTCAGGCTTGGAGAGTATGACCTCGCGGCTGGGAGAGTGGGAGCTGGACCTGGACATC 720  
Db 221 ValArgLeuGlyGlyTyArgLeuArgTrpGluLysTrpGluLeuAspLeuAspIle 240  
QY 721 AAGAGGTCTTCTCCACCCCACTACAGCAAGAGCACCACCGACATGACATCGCACTG 780  
Db 241 LysGluValPheValHisProAsnTyrlsSerLysSerThrThrAspAsnAspIleAlaLeu 260  
QY 781 CTGCACTTGGCCCGCAGCCGCGCCCTCTCCAGACCATAGTGCCTATGCTCCCGGAC 840  
Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 280  
QY 841 AGCGGCTTGCAGAGCGGAGCTCAATCAGCGCGCCGAGGAGACCTCTGAGCGGCTGG 900  
Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
QY 901 GGCTACCACAGCAGCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
Db 301 GlyTyrlsSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
QY 961 ATCAAGATTCCTGCTCCCGCACAATGATGTCAGGAGGTCATGAGCAACATGCTGTCT 1020  
Db 321 IleLysIleProValValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
QY 1021 GACAACATGCTGTGCTGGGCAATCTCGGGGACCGGACGAGATGCTGGAGGCGGACAGT 1080  
Db 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
QY 1081 GGGGGCCCATGGTCCCTCTCCAGCGCAGCTGGTTCCTGGTGGGCTGTGAGCTGG 1140  
Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
QY 1141 GGTGAGGCTGTGGGCTCTTCAACTACGGCGTTTACACCAAGTACAGCGCTGTCTC 1200  
Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTyrglyValTyThrLysValSerArgTyLeu 400  
QY 1201 GACTGATCCATGGGCACATCAGACAAAGAAAGCCGCCCAAGAAG 1245  
Db 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415

## RESULT 6

US-10-182-263-3  
; Sequence 3, Application US/10182263  
; Publication No. US20030022354A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerlitz, Bruce E  
; APPLICANT: Jones, Bryan E  
; APPLICANT: Grinnell, Brian W  
; TITLE OF INVENTION: PROTEIN C DERIVATIVES  
; FILE REFERENCE: X-13611  
; CURRENT APPLICATION NUMBER: US/10/182,263  
; CURRENT FILING DATE: 2002-07-22  
; PRIOR APPLICATION NUMBER: 60/1891948

;; PRIOR FILING DATE: 2002-02-11  
;; PRIOR APPLICATION NUMBER: 60/189199  
;; PRIOR FILING DATE: 2000-03-14  
;; NUMBER OF SEQ ID NOS: 12  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 3  
;; LENGTH: 419  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-182-263-3

Alignment Scores:  
Pred. No.: 1,46e-125 Length: 419  
Score: 2264.00 Matches: 409  
Percent Similarity: 99.04% Conservative: 2  
Best Local Similarity: 98.55% Mismatches: 4  
Query Match: 96.88% Indels: 0  
DB: Gaps: 0

US-09-763-153-1 (1-1245) x US-10-182-263-3 (1-419)

1 GCCAACTCTCTCTGGAGGAGTCCTGCACAGACGCTGGAGCGGAGTGCATAGAGAG 60  
1 AlaAsnSerPheLeuGluGluLeuArgHisGlySerLeuGluArgGluCysIleGlu 20  
61 ATCTGTGACTTCGAGAGCGCAAGGAATTTCCAAATGTTGGATGACACACTGGCTTC 120  
21 IleCysAspPheGluGluAlaLysGluIlePheGluAspValAspThrLeuAlaPhe 40  
121 TGGCTCAAGCAGCTCCAGCGTGCACAGTGTCTTGGCTTGGCTGGAGCACCCTGGCGC 180  
41 TrpSerLysHisValAspGlyaspGlnCysLeuValLeuProLeuGluHisProCysAla 60  
181 AGCTGTGCTGCGGCGCAGCGCATGTCATCGACGGCATCGGAGCTTCACTGCGACTGC 240  
61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
241 CGCAGCGCTGGAGGCGGCTCTCCAGCGGAGTGGCTTCTCAATTTGCTGCGCTG 300  
81 ArgSerGlyTrpGluArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
301 GACAAACGCGCTGCACGATTAATCTGCTAGAGAGTGGCTGGCGGCTGTAGCTGT 360  
101 AspAsnGlyGlyCysThrHisTyrCysLeuGluGluValGlyTrpArgArgCysSerCys 120  
361 CGCCTGGCTACAGCTGGGGGAGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
121 AlaProGlyTyrLysLeuGlyaspGlnCysHisProAlaValLysPhePro 140  
421 TGTGGAGGCGCTGGAAGCGGATGGAGAGAGCGCAGTCACTGAAACGACACAGAA 480  
141 CysGlyArgProTyrLysArgMetGluLysLysArgSerHisLeuLysArgAspThrGlu 160  
481 GACCAAGAACACCAAGTAGATCCGCGCTCATTCATGGGAGATGACACGCGGGAGAC 540  
161 AspGlnGluAspGlnValPheProArgLeuIleGlyLysMetThrArgArgGlyasp 180  
541 AGCCCTGGCAGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
181 SerProTyrGlnValValLeuLeuAspSerLysLysSerAlaCysGlyAlaValLeu 200  
601 ATCCACCCCTCTGGGTGCTGACAGCGGCCCTGATGATGATGATGATGATGATGATGAT 660  
201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLysLeu 220  
661 GTCAGCTTGGAGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
221 ValArgLeuGlyLysTrpValPheArgArgTrpGluLysTrpGluLeuAspPheAla 240  
721 AAGGAGGTCTTGTCCACCACTACAGCAAGAGACACCGCAATGATGATGATGATGATGAT 780  
241 LysGluValPheValHisProAsnTyrSerLysSerThrThrAspAsnAspIleAlaLeu 260

781 CTGCACCTGGCCAGCCCGCCACCCCTCTCTCAGACCATAGTGCCCATCTGCTCCGGAC 840  
261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProIleCysLeuProAsp 280  
841 AGCGGCTTGGAGAGCGGAGCTCAATCAGCGCGGCGGAGGAGCCTGCTGAGCGGCTGG 900  
281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
901 GGCTACACACAGCGCGGAG 960  
301 GlyTyrHisSerSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
961 ATCAAGATTCCTGGTCCCGCCACCAATGATGAGTGCAGGAGTGCATGACACATGGTGTCT 1020  
321 IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
1021 GAGAACATGCTGTGTGGCGGATCTCTGGGAGCGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
1081 GGGGGCGCCATGCTGCTCTCTCCACGGCACCTGGTTCCTGGTGGCGCTGGTGGCTGG 1140  
361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
1141 GGTGAGGCTGTGGCTCTCTCACAACCTACCGCGCTTACACCAAGTCAGCGCTACCTC 1200  
381 GlyGluGlyCysGlyLeuLeuHisAsnTyrGlyValTyrThrLysValSerArgTyrLeu 400  
1201 GACTGGATCCAGTGGCAGCATCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245  
401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415

RESULT 7

US-10-182-263-6  
; Sequence 6, Application US/10182263  
; Publication No. US20030022354A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerlitz, Bruce E  
; APPLICANT: Jones, Bryan E  
; APPLICANT: Grinnell, Brian W  
; TITLE OF INVENTION: PROTEIN C DERIVATIVES  
; FILE REFERENCE: X-13611  
; CURRENT APPLICATION NUMBER: US/10/182,263  
; CURRENT FILING DATE: 2002-07-22  
; PRIOR FILING DATE: 2002-02-11  
; PRIOR APPLICATION NUMBER: 60/189199  
; PRIOR FILING DATE: 2000-03-14  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 419  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-182-263-6

Alignment Scores:  
Pred. No.: 1,91e-125 Length: 419  
Score: 2262.00 Matches: 409  
Percent Similarity: 99.04% Conservative: 2  
Best Local Similarity: 98.55% Mismatches: 4  
Query Match: 96.79% Indels: 0  
DB: Gaps: 0

US-09-763-153-1 (1-1245) x US-10-182-263-6 (1-419)

1 GCCAACTCTCTCTGGAGGAGTCCTGCACAGACGCTGGAGCGGAGTGCATAGAGAG 60  
1 AlaAsnSerPheLeuGluGluLeuArgHisGlySerLeuGluArgGluCysIleGlu 20  
61 ATCTGTGACTTCGAGAGCGCAAGGAATTTCCAAATGTTGGATGACACACTGGCTTC 120  
21 IleCysAspPheGluGluAlaLysGluIlePheGluAspValAspThrLeuAlaPhe 40



QY 121 TGGTCCAAAGCAGCTGACGAGTACCAGTGGCTTGGTCTTGGCTTGGAGCACCCTGGCC 180  
Db 41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60  
QY 181 AGCTGTGTGGGGCAGCGACGTCATCGACGCGATCGGCGCTTCAGCTGCGACTGC 240  
Db 61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
QY 241 CGCAGCGGTGGAGGGCGCTTCTGCCAGCGCGAGGTGAGCTTCCCAATTCGCTGCTG 300  
Db 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
QY 301 GACAAAGCGGTGCACGATTAATCTGCTAGAGAGGTGGCTGGCGGCTGTAGCTGT 360  
Db 101 AspAsnGlyGlyCysThrHisThrCysLeuGluValGlyTrpArgArgCysSerCys 120  
QY 361 GCGCTGTGTACAGCTGGGGAGCAGCTTCTGCAGTGTCCACCGCAGTGAAGTTCCT 420  
Db 121 AlaProGlyTrpLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 140  
QY 421 TGTGGAGCGCTCGAAGCGGATGGAGAGAGCGACCTACCTGAAACGACAGACAGAA 480  
Db 141 CysGlyArgProTrpLysArgMetGluLysArgSerHisLeuLysArgAspThrGlu 160  
QY 481 GACCAAGAACCAAGTAGATCCGCGCTCATTTGATGGAGAGATGACGAGCGGGGAGAC 540  
Db 161 AspGlnGluAspGlnValPheProArgLeuIleLysGlyLysMetThrArgArgGlyAsp 180  
QY 541 AGCCCTGGAGGTGGTCTGCTGGACTCAAGAGAGAGTGGCGCTGGCGGCGAGTC 600  
Db 181 SerProTrpGlnValValLeuLeuAspSerLysLysLysLeuAlaCysGlyAlaValLeu 200  
QY 601 ATCCACCCCTCTGGGTGCTGACAGCGCGCCACTGTGATGATGAGTCCCAAGAGCTCTT 560  
Db 201 IleHisProSerTrpValLeuThrAlaAlaHisCysMetAspGluSerLysLeuLeu 220  
QY 661 GTCAGGCTTGAGAGTAGTACCTCGCGCTGGGAGAGTGGGAGCTGGACCTGGACATC 720  
Db 221 ValArgLeuGlyGluTrpAspLeuArgArgTrpGluLysTrpLeuAspLeuAspIle 240  
QY 721 AAGCAGCTCTGCGCCACCCCACTACAGCAAGAGACACACGCAATGACATCGCAGTC 780  
Db 241 LysGluValPheValHisProAsnTrpSerLysSerThrIleValProLysLeuProAsp 260  
QY 781 CTGACCTGGCCCGCCAGCCCTCTCGCAGACCATAGTGCCTATCTGCTCCCGGAC 840  
Db 261 LeuHisLeuAlaGlnProAlaThrLeuSerGlnThrIleValProLysLeuProAsp 280  
QY 841 AGCGGCTTCGAGCGCGAGCTCAATCAGCGCGCGCCAGGACCTCGTGACGGGCTGG 900  
Db 281 SerGlyLeuAlaGluArgGluLeuAsnGlnAlaGlyGlnGluThrLeuValThrGlyTrp 300  
QY 901 GGCTACACAGCAGCGAGAGAGGAGGCGCAAGAGAAACCGCACCTTCTGCTCAACTC 960  
Db 301 GlyTrpHisSerArgGluLysGluAlaLysArgAsnArgThrPheValLeuAsnPhe 320  
QY 961 ATCAAGATCCCTGTGTCGCCCAATAGTGCAGCGAGTGCATGAGCAACATGTGTCT 1020  
Db 321 IleLysIleProValProHisAsnGluCysSerGluValMetSerAsnMetValSer 340  
QY 1021 GAGAACATGCTGTGCGGGCATCTCGGGGACCGGCGAGAGTCCGAGGGCGACAGT 1080  
Db 341 GluAsnMetLeuCysAlaGlyIleLeuGlyAspArgGlnAspAlaCysGluGlyAspSer 360  
QY 1081 GGGGGGCCATGCTGCTCTTCCAGGCGACCTGTGCTGCTGGTGGGCTGTGAGCTGG 1140  
Db 361 GlyGlyProMetValAlaSerPheHisGlyThrTrpPheLeuValGlyLeuValSerTrp 380  
QY 1141 GGTGAGGCTGTGGGCTCTCACAACACTACGGCGTTTACACAAAGTCAGCGCTACTCTC 1200  
Db 381 GlyGluGlyCysGlyLeuLeuHisAsnTrpGlyValTrpThrLysValSerArgTrpLeu 400

QY 1201 GACTGTATCCATGGCAGATCAGACAAAGGAAGCCCGCCAGAAAG 1245  
Db 401 AspTrpIleHisGlyHisIleArgAspLysGluAlaProGlnLys 415

## RESULT 8

US-10-182-263-4  
; Sequence 4, Application US/10182263  
; Publication No. US20030022354A1  
; GENERAL INFORMATION:  
; APPLICANT: Jones, Bryan E  
; APPLICANT: Gerlitz, Bruce E  
; APPLICANT: Grimell, Brian W  
; TITLE OF INVENTION: PROTEIN C DERIVATIVES  
; FILE REFERENCE: X-13611  
; CURRENT APPLICATION NUMBER: US/10/182,263  
; CURRENT FILING DATE: 2002-07-22  
; PRIOR APPLICATION NUMBER: 60/181948  
; PRIOR FILING DATE: 2002-02-11  
; PRIOR APPLICATION NUMBER: 60/189199  
; PRIOR FILING DATE: 2000-03-14  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 419  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-182-263-4  
Alignment Scores:  
Pred. No.: 2,5e-125 Length: 419  
Score: 2260.00 Matches: 408  
Percent Similarity: 99.04% Conservative: 3  
Best Local Similarity: 98.31% Mismatches: 4  
Query Match: 96.71% Indels: 0  
DB: Gaps: 0

US-09-763-153-1 (1-1245) x US-10-182-263-4 (1-419)

QY 1 GCCAACTCTCTTCTGGAGGAGCTCCCTCACAGCAGCGCTGGAGCGGAGTGCATAGAGAG 60  
Db 1 AlaAsnSerPheLeuGluGluLeuArgHisGlySerLeuGluArgGluCysIleGluGlu 20  
QY 61 ATCTGTCACTTCGAGGAGGCGCAAGGAAATTTCCAAATGTGGATGACACACTGGCTTC 120  
Db 21 IleCysAspPheGluGluAlaLysGluIlePheGluAspValAspAspThrLeuAlaPhe 40  
QY 121 TGTCTCAAGCAGCTCGACGCTGACAGTGTCTGCTTGGCTTGGAGCAGCGCTGGCGC 180  
Db 41 TrpSerLysHisValAspGlyAspGlnCysLeuValLeuProLeuGluHisProCysAla 60  
QY 181 AGCTGTGTGCGGGCAGCGCAGCTGCATCGCAGCGGATCGGAGCTTCAGCTGCGACTGC 240  
Db 61 SerLeuCysCysGlyHisGlyThrCysIleAspGlyIleGlySerPheSerCysAspCys 80  
QY 241 CGCAGCGGTGGAGGGCGCTTCTGCCAGCGCGAGGTGAGCTTCCCAATTCGCTGCTG 300  
Db 81 ArgSerGlyTrpGluGlyArgPheCysGlnArgGluValSerPheLeuAsnCysSerLeu 100  
QY 301 GACAAAGCGGTGCACGATTAATCTGCTAGAGAGGTGGCTGGCGGCTGTAGCTGT 360  
Db 101 AspAsnGlyGlyCysThrHisThrCysLeuGluValGlyTrpArgArgCysSerCys 120  
QY 361 GCGCTGTGTACAGCTGGGGAGCAGCTTCTGCAGTGTCCACCGCAGTGAAGTTCCT 420  
Db 121 AlaProGlyTrpLysLeuGlyAspAspLeuLeuGlnCysHisProAlaValLysPhePro 140  
QY 421 TGTGGAGCGCTCGAAGCGGATGGAGAGAGCGACCTACCTGAAACGACAGACAGAA 480  
Db 141 CysGlyArgProTrpLysArgMetGluLysArgSerHisLeuLysArgAspThrGlu 160  
QY 481 GACCAAGAACCAAGTAGATCCGCGCTCATTTGATGGAGAGATGACGAGCGGGGAGAC 540  
Db 161 AspGlnGluAspGlnValPheProArgLeuIleLysGlyLysMetThrArgArgGlyAsp 180

Score: 736.00		Matches: 150	
Percent Similarity: 52.36%		Conservative: 72	
Best Local Similarity: 35.38%		Mismatch: 156	
Query Match: 31.49%		Indels: 46	
DB:		Gaps: 10	
US-09-763-153-1 (1-1245) x US-10-132-829-5 (1-461)			
QY	13	CTGAGAGAGCTCCGTCACACGACGCTGGAGCGGAGTGCATAGAGGAGATCTGTGACTTC	72
Db	52	LeuGluGluPheValGlnGlyAsnLeuGluArgGlyCysMetGluGluLysCysSerPhe	71
QY	73	GAGGAGCGCAAGGAATTTCCAAATGTGGATGACACACATGGCCCTTGTGGTCCAAAGAC	132
Db	72	GluGluAlaArgGluValPheGluAsnThrGluArgThrThrGluPheTrpLysGlnTyr	91
QY	133	GTGACGGGTGACCAAGTCTGTGCTTCCCTTGGAGACCCCTGGCCGAGCTGTGCTGC	192
Db	92	ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu	103
QY	193	GGGACGCGCAGCTGCATCGAGCGGATCGGCGCTTCAGCTGCGACTGCGCGAGCGGCTGG	252
Db	104	AsnGlyGlySerCysLysAspAspIleAsnSerTyrGluCysTrpCysProPheGlyPhe	123
QY	253	CAGGCGCGCTTCTGCCAGCGGAGTGTGCTTCTCAATTTGCTGCTGGTGCACAAACGCGGC	312
Db	124	GluGlyLysAsnCysGluLeuAspValThr-----CysAsnIleLysAsnGlyArg	140
QY	313	TGCAGCGATTACTGCCTAGAGGAGTGGCTGGCGG---CGCTGTAGCTGTGCGCTGGC	369
Db	141	CysGluGlnPheCysLysAsnSerAlaAspAsnLysValValCysSerCysThrGluGly	160
QY	370	TACAAAGCTGGGGGAGCGACCTCTGCTGCTACCGCGAGTGAAGTTCCTTGTGGGAGG	429
Db	161	TyrArgLeuAlaGluAsnGlnLysSerCysGluProAlaValProPheProCysGlyArg	180
QY	430	CCCTGGAGCGGATGGAGAAAGCGGAGTCCCTGAAACGA-----	471
Db	181	ValSerValSerGlnThrSerLysLeuThrArgAlaGluThrValPheProAspValAsp	200
QY	472	-----GACACAGAAAGACCAAGAACCAAGTAGAT-----	501
Db	201	TyrValAsnSerThrGluAlaGluThrIleLeuAspAsnIleThrGlnSerThrGlnSer	220
QY	502	-----CCCGGCTCATGATGGGAAGATGACCGCGGGGAGAGACAGCCCTGG	549
Db	221	PheAsnAspPheThrArgValValGlyGlyGluAspAlaLysProGlyGlnPheProTrp	240
QY	550	CAGGTGGTCTCTGCTGACTCAAGAAAGAGTGGCTGCGGCGGAGTGCATCCACCCC	609
Db	241	GlnValVal---LeuAsnGlyLysValAspAlaPheCysGlyGlySerIleValAsnGlu	259
QY	610	TCTGTGGTCTGACAGCGGCCCTGCTGATGATGATGATGATGATGATGATGATGATG	669
Db	260	LysTrpIleValThrAlaAlaHisCysValGluThrGlyValLysIleThrValAla	279
QY	670	GGAGATATGACCTGGCGCTGGAGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT	729
Db	280	GlyGluHisAsnIleGluThrGluHisThrGluGlnLysArgAsnValIleArgIle	299
QY	730	TTCGTCTCCCACTACAGCAAGACGACCCACC-----GACAATGACATCGCTGCTG	783
Db	300	IleProHisHisAsnTyrAsnAlaAlaIleAsnLysTyrAsnHisAspIleAlaLeuLeu	319
QY	784	CACCTGGCGCGCGCCACCTCTCGCAGACCATAGTGGCCCATCTGCTCCCGGACAGC	843
Db	320	GluLeuAspGluProLeuValLeuAsnSerTyrValThrProIleCysIleAlaAspLys	339
QY	844	GGCTTGCAGAGCGGAGCTCAATCAGCGCGGCGGAGGAGACCTCGTGCAGCGGTGGGC	903
Db	340	GluTyrThrAsnIlePheLeuLysPheGly-----SerGlyTyrValSerGlyTrpGly	357
QY	904	-----TACCACAGCAGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	957

US-10-132-829-5

sequence 5, Application US/10132829

Publication No. US20030044982A1

GENERAL INFORMATION:

APPLICANT: Chien, Kenneth R

APPLICANT: Hoshijima, Masahiko

TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor VII

FILE OF INVENTION: with vesicle vector

FILE REFERENCE: 6627-PAL170

CURRENT APPLICATION NUMBER: US/10/132,829

CURRENT FILING DATE: 2002-04-25

PRIOR APPLICATION NUMBER: 60/286,314

PRIOR FILING DATE: 2001-04-25

NUMBER OF SEQ ID NOS: 5

SOFTWARE: PatentIn version 3.1

SEQ ID NO 5

LENGTH: 461

TYPE: PRT

ORGANISM: Homo sapiens

US-10-132-829-5

Alignment Scores: 5.21e-36 Length: 461

Pred. No.:

US-10-132-829-5  
Sequence 5, Application US/10132829  
Publication No. US20030044982A1  
GENERAL INFORMATION:  
APPLICANT: Chien, Kenneth R  
APPLICANT: Hoshijima, Masahiko  
TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor VII  
TITLE OF INVENTION: with vesicle vector  
FILE REFERENCE: 6627-PALL70  
CURRENT APPLICATION NUMBER: US/10/132,829  
CURRENT FILING DATE: 2002-04-25  
PRIOR APPLICATION NUMBER: 60/286,314  
PRIOR FILING DATE: 2001-04-25  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5  
LENGTH: 461  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-132-829-5  
Alignment Scores: 5.21e-36 Length: 461  
Pred. No.:

```

Db      358 ArgValPheHisLysGlyArgSer-----AlaLeuValLeuGln   370
          ::::||| ||| |||||
QY      958 TTCATCAAGATTCCGTGGCGGCACAATGAGTCACGAGGTCAATGAGCAACATGGTG   1017
          ::::::::::::::|||
Db      371 TyrLeuArgValProLeuValAspArgAlaThrCysLeuArgSerThrLysPheThrIle   390
          ::::::::::::::|||
QY     1018 TCTCAGAACATGCTGTGTGGGGCATCTCCGGGACCGCACGAGTGCCTCGAGGGCGAC   1077
          ||||| |||||||
Db      391 TyrAsnAsnMetPheCysAlaGlyPheHisGluGlyArgAspSerCysGlnGlyAsp   410
          ||||| |||||||
QY     1078 AGTGGGGGCCCATGGTGGCGCTCTTCCACGGCACCTGTTCCCTGGTGGCGCTGGTGAGC   1137
          ||||| ||||| ||||| |||||
Db      411 SerGlyGlyProHisValThrGluValGluGlyThrSerPheLeuThrGlyIleIleSer   430
          ::::::::::::::|||
          1138 TGGGGTAGGGCGCTGGGGCTCTCTCACAACTACGGCGTTTTACACCAGATCAGCCGCTAC   1197
          ||||| ::::
          431 TrpGlyglucLucyasaAlaMetLysGlyLystyrGlyIleTyrrThrLysValSerArgtyr   450
          ||||| ::::::|||
QY     1198 CTCGACTGGATC 1209
Db      451 ValasnIrpile 454

RESULT 10
US-09-884-901-3
; Sequence 3, Application US/09884501
; Patent No. US2002007698A1
; GENERAL INFORMATION:
; APPLICANT: Miao, Carol
; APPLICANT: Kay, Mark
; TITLE OF INVENTION: Liver-Specific Gene Expression Cassettes, and Methods
; FILE REFERENCE: UOFW-1-17396
; CURRENT APPLICATION NUMBER: US/09/884,901
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/212,902
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 3
; LENGTH: 461
; TYPE: PRT
; ORGANISM: HomoSapien
US-09-884-901-3
```

QY	313	TCGACGCATTACTGCCTTAGGAGGTTGGCTGCGCG---CGCTGTAGCTGTGCGCCTGGCG	369
Db	141	CysGluInPheCysLysAsnSerAlaAspAsnLysValValCysSerCysThrGluGly	160
QY	370	TACAAGCTGGGGGACGACCTCCTGCACTGTCACCCGCACTGAAGTTCCTTGTGGGAGG	429
Db	161	TyrArgLeuAlaGluAsnGlnLysSerCysGluProAlaValProPheProCysGlyArg	180
QY	430	CCCTTGGAGCGGATGGAGAGAAGCGCGATCACTGAAACGA-----	471
Db	181	ValSerValSerGlnThrSerLysLeuThrArgAlaGluAlaValPheProAspValAsp	200
QY	472	-----GACACGAGAACCAAGAGACCAAGTAGAT-----	501
Db	201	TyrValAsnSerThrGluAlaGluThrIleLeuAspAsnIleThrGlnSerThrGlnSer	220
QY	502	-----CGCGGCTATTGATGGAGAGATGACGAGGGGGAGACAGCCCTGG	549
Db	221	PheAsnAspPheThrArgValValGlyGluAspAlaLysProGlyGlnPheProThr	240
QY	550	CAGTGTCTCTGTGACTCAAAAGAAAGCTGGCCTCGGGGCGAGTCTATCCACCC	609
Db	241	GlnValVal--LeuAsnGlyLysValAspAlaPheCysGlyCysIleValAsnGlu	259
QY	610	TCCTGGTGTGACAGGGCCCACTCCATGGATGAGTCCAGAAGCTCTTGTAGGCTT	669
Db	260	LysTrpIleValThrAlaAlaHisCysValGluThrGlyValLysIleThrValValAla	279
QY	670	GGAGAGTATCACTCGCGGCTGGGAGAGTGGGAGCTGGACCTGGACATCAAGGAGTC	729
Db	280	GlyGluHisAsnIleGluGluThrGluHisThrGluGlnLysArgAsnValIleArgIle	299
QY	730	TTGCTGCCACCCCACTACAGCAGACCAACC-----GACAATGACATCGCACTGTG	783
Db	300	IleProHisHisAsnTyrAsnAlaAlaIleAsnLysTyrAsnHisAspIleAlaLeuLeu	319
QY	784	CACTCGGCCACGCCCCCTCTCGCAGACCATAGTCCCATCTGCTCCCGGACAGC	843
Db	320	GluLeuAspGluProLeuValLeuAsnSerTyrValThrProIleCysIleAlaAspLys	339
QY	844	GGCTTGCAGAGCGGAGTCAATCAGCGCGCCAGGAGACCCCTGCTGACGGGTGGGGC	903
Db	340	GluTyrThrAsnIlePheLeuLysPheGly-----SerGlyTyrValSerGlyTrpGly	357
QY	904	-----TACCACAGCAGCGGAGAGAGGAGGCGCAAGAACCCGACCTTGTCTCTCAAC	957
Db	358	ArgValPheHisGlyArgSer-----AlaLeuValLeuGln	370
QY	958	TTTCATCAAGATTCCCGTGGTCCCGCACAAATGAGTGCCAGCGAGGTCAATGAGCAACATGGT	1017
Db	371	TyrLeuArgValProLeuValAspArgAlaThrCysLeuArgSerThrLysPheThrIle	390
QY	1018	TCTGAACATCTGTGTGGGCATCTCTCGGGGACCGGACAGGATGCTCGGAGGGCGAC	1077
Db	391	TyrAsnAsnMetPheCysAlaGlyPheHisGluGlyArgAspSerCysGlnGlyAsp	410
QY	1078	AGTGGGGGCCCATCGTCCCTCCCTCCACCGCACCTCGTTTCCTTGGTGGCCCTGGTGA	1137
Db	411	SerGlyGlyProHisValThrGluValGluGlyThrSerPheLeuThrGlyIleIleSer	430
QY	1138	TGGGTGAGGGCTGTGGCTCTCTTACACACTACCGGTTCACACCAAGTCACGCCCTAC	1197
Db	431	TrpGlyGluGluCysAlaMetLysGlyLysTyrGlyIleTyrThrLysValSerArgTyr	450
QY	1198	CTCAGACTGATC	1209
Db	451	ValAsnTrpIle	454

RESIST 11

RESULT II  
US-09-118-748-2  
; Sequence 2, Application US/09118748A  
; Patent No. US20020031799A1

```

; GENERAL INFORMATION:
; APPLICANT: Staiford, Darrel W.
; APPLICANT: Chang, JinLi
; TITLE OF INVENTION: Factor Ix Antihemophilic Factor with Increased Clotting
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 5470-183
; CURRENT APPLICATION NUMBER: US/09/118,748A
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/053,571
; EARLIER FILING DATE: 1997-07-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-118-748-2

Alignment Scores:
Seq. No.: 5,94e-36 Length: 415
Percent Identity: 735.00 Matches: 150
Percent Similarity: 52.36% Conservative: 72
Query Match: 31.45% Mismatches: 156
DB: 10 Gaps: 10

US-09-763-153-1 (1-1245) x US-09-118-748-2 (1-415)
QY 13 CTGAGGAGGCTCGCTCACAGCAGCCTGGAGCGGGAGTGCATAGAGGAGATCTGTGACTTC 72
DB 6 LeuGluGluPheValGlnGlyAsnLeuGluArgGluCysMetGluGlyCysSerPhe 25
QY 73 GAGGAGGCCAAGAAATTTTCCAAAATGTGGATGACACACTGGCCTCTGTGTCCAAAGCAC 132
DB 26 GluGluAlaArgGluValPheGluAsnThrGluArgThrGluPheTrpLysGlnTrp 45
QY 133 GTGACGGGTGACAGCGCTGTGTCTGTGCTTGGAGCACCCTGGCCAGCCTGTGTGTGC 192
DB 46 ValAspGlyAspGlnCysGluSerAsnPro-----CysLeu 57
QY 193 GGGCAGCGGACGTCATCGACGCGATCGGCAGCTTCAGCTGGAGTGGCGGCGGCTGG 252
DB 58 AsnGlyGlySerCysLysAspIleAsnSerTyrGluCysTrpCysProPheGlyPhe 77
QY 253 GAGGGCGCGCTTGTCCAGCGCGAGGTGAGCTTCCTCAATTGCTCGCTGGACACGCGCGC 312
DB 78 GluGlyLysAsnCysGluLeuAspValThr-----CysAsnIleLysAsnGlyArg 94
QY 313 TGCACGCATTACTGCTCCCTAGAGGAGGTGGCTGGCGG---CGCTGTAGCTGTGCGCTGGC 369
DB 95 CysGluGlnPheCysLysAsnSerAlaAspAsnLysValValCysSerCysThrGluGly 114
QY 370 TACAAGCTGGGGGACGACCTTCCTGCAGTGTACCCCGCAGTGAAGTTCCTTTGTGGGAGG 429
DB 115 TyrArgLeuAlaGluAsnGlnLysSerCysGluProAlaValProPheProCysGlyArg 134
QY 430 CCCTGGAGCGGATGGAGAGAGCGGACGTCACCTGAACGA----- 471
DB 135 ValSerValSerGlnThrSerLysLeuThrArgAlaGluThrValPheProaspValasp 154
QY 472 -----GACACAGAACACCAAGAGACCAAGTAGAT----- 501
DB 155 TyrValAsnSerThrGluAlaGluThrIleLeuAspAsnIleThrGlnSerThrGlnSer 174
QY 502 -----CCGCGGCTCATTTGATGGGAAGATGACACAGGCGGGGACAGCCCTGG 549
DB 175 PheAsnAspPheThrArgValValGlyGlyAspAlaLysProGlyGlnPheProTrp 194
QY 550 CAGGTGGTCTGCTGGACTCAAGAAAGAGAGCTGGCGTGGCGGACGTGTCTCATCACCC 609
DB 195 GlnValVal-----LeuAsnGlyLysValAspAlaPheCysGlyGlySerIleValAsnGlu 213
QY 610 TCCTGGGTGCTGACAGCGGGGCCACTGCTGGATGGATGAGTCCAAAGAGCTTCCTGTGACGGCT 669

```

Best Local Similarity:	36.64%	Mismatches:	157
Query Match:	31.07%	Indels:	36
DB:	9	Gaps:	10

US-09-763-153-1 (1-1245) x US-10-109-498-1 (1-406)

QY	1	GC	CAACTCCTTCTCGAGGAGCTCCGTACAGCAGCGCTCGAGCGGAGTGCATAGGAGG	60
Db	1	Ala	AsnAlaPheLeu*****LeuArgProGlySerLeu***Arg***CysLys*****	20
QY	61	ATC	TGTGACTTCGAGGAGGCCAAGAAATTTCCAAATGTGATGACACACTGGCCTTC	120
Db	21	Gln	CysSerPhe*****AlaArg***IlePheLysAspAla***ArgThrLysLeuPhe	40
QY	121	TGG	TCCAAGCAGCTCGACGGTGACAGTCTGGTCTGCGCTTGAGCACCCTGCGCC	180
Db	41	Trp	IleSerTyrSerAspGlyAspGlnCys-----AlaSer	52
QY	181	AGC	TGTGCTCGCGGCACGCGATGCGATGCGCGATCGGAGTTCAGCTCGCAGTCG	240
Db	53	Ser	ProCysGlnAsnGlyGlySerCysLysAspGlnSerTyrIleCysPheCys	72
QY	241	CGC	AGCGCTGGGAGGCCCTCTCCAG---CGCGAGTGAGCTTCTCAATTGCTCG	297
Db	73	Leu	ProAlaPheGluGlyArgAsnCysGluThrHisLysAspGlnLeuIleCysVal	92
QY	298	CTG	GACAAACGCGGCTGCACGCAATTACTCGCTAGAGAGGTGGTGGCGGCG---NGT	354
Db	93	Asn	GluAsnGlyCysGluGlnTyrCysSerAspHisThrGlyThrLysArgSerCys	112
QY	355	AGC	TGTGCGCTGGCTACAGCTGGGGAGACCTCTCTGCTAGTCACCCGACGTGAAG	414
Db	113	Arg	CysHisGluGlyTyrSerLeuLeuAlaAspGlyValSerCysThrProThrValGlu	132
QY	415	TTCC	TGTGGGAGGCCCTGGAAGCGATGAGAGAAGCGCAGTCACCTGAACGAGAC	474
Db	133	Tyr	ProCysGlyLys---IleProIleGluGluLysArgAsnAla-----	146
QY	475	AC	GAGACCCAGAAGCAACGATAGATCGCGGCTCATTTGATGGGAAGATCACCAGCGG	534
Db	147	-----	SerLysProGlnGlyArgIleValGlyGlyLysValCysProLys	161
QY	535	GG	ACAGACGCCCTGGCAGCTGTCTCTGGACTCAAGACAAGCACTGGCTGGCGGCA	594
Db	162	Gly	CysProTrpGlnValLeuLeuLeuValAsnGlyAlaGlnLeu---CysGlyGly	180
QY	595	GT	CTCATCACCCCTCTCGGTGTGACAGCGGCCACTGCATGATGATGCCAAG---	651
Db	181	Thr	IleAsnThrIleTrpValValSerAlaAlaHisCysPheAspLysIleLysAsn	200
QY	652	-----	AACTCTTCTGAGGATGATGACCTCGCGGCTGGGAGAAAGTGGGAG	705
Db	201	Trp	ArgAsnLeuIleAlaValLeuGlyGluHisAspLeuSerGluHisAspGlyAspGlu	220
QY	706	CTG	ACCTCGACCAAGGAGTCTTCGTCCACCCCACTACACGACGAGCACCACCGAC	765
Db	221	Gln	SerArgValAlaGlnValIleIleProSerThrTyrValProGlyThrThrAsn	240
QY	766	AAT	GATCGCACTGCTGCACTGGCCACGCGCGCCACCTCTCGCAGACCATAGTCCGC	825
Db	241	His	AspIleAlaLeuLeuArgLeuHisGlnProValValLeuThrAspHisValPro	260
QY	826	ATC	TGCTCCCGACAGCGGCTTTCAGAGCGGAGTCAATCAGGCGGCCAGGAGACC	885
Db	261	Leu	CysLeuProGluArgThrPheSerGluArgThrLeuAlaPheVal---ArgPheSer	279
QY	886	CTC	GTGACGGGTGGGCTACACAGCAGCCGAGAGAGGAGGCGCAAGAAACCGCAC	945
Db	280	Leu	ValSerGlyTrpGlyGlnLeuLeuAspArgGlyAlaThrAla-----	294
QY	946	TT	CGCTCAACTTTCATCAAGATTCCTGGTGTCCCGCACATGATGATCGCAGGTCATG	1005
Db	295	Leu	GluLeuMetValLeuValAsnValProArgLeuMetThrGlnAspCysLeuGlnSer	314



1

;; PRIOR APPLICATION NUMBER: 60/077791  
;; PRIOR FILING DATE: 1998-03-12  
;; PRIOR APPLICATION NUMBER: 60/078004  
;; PRIOR FILING DATE: 1998-03-13  
;; PRIOR APPLICATION NUMBER: 60/078886  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078936  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078939  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/079656  
;; PRIOR FILING DATE: 1998-03-26  
;; PRIOR APPLICATION NUMBER: 60/079664  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079689  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079663  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079728  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079786  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079920  
;; PRIOR FILING DATE: 1998-03-30  
;; PRIOR APPLICATION NUMBER: 60/079923  
;; PRIOR FILING DATE: 1998-03-30  
;; PRIOR APPLICATION NUMBER: 60/080105  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080107  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080165  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080194  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080327  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080328  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080333  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080334  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/081070  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081049  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081071  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081195  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081955  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081819  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081952  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081838  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082568  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082569  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082704







Pred. No.:	2,07e-20	Length:	802
Score:	470.00	Matches:	124
Percent Similarity:	46.93%	Conservative:	52
Best Local Similarity:	33.07%	Mismatches:	145
Query Match:	20.11%	Indels:	54
	9	Gaps:	18

DB:

US-09-763-153-1 (1-1245) x US-09-978-192A-169 (1-802)

QY	172	CGGTGGCCGAGC	---CTGTGCTGGGACAGCGACGTGCATC-----GAGCGC	216
DB	448	PrOCysProGlyGluPheLeuCysSerValAsnGlyLeuCysValProAlaCysAspGly	467	
QY	217	ATCGGCAGTTCAGCTGCGATCGCGAGCGCGCTGGGAGGCGCGCTTCCAGCGCGAG	276	
		:::		:::
DB	468	ValLys-----AspCysProAsnGlyLeuAspGluArgAsnCysValCysArg	483	
DB	277	GTGAGTTCCTCAATTCGTGCTGGACAAACGCGCGCTGCACGCATTTACTGCTTAGAGGAG	336	
		:::		:::
DB	484	AlaThrPhe-----GlnCysLysGluAspSer-----ThrCysIleSerLeuProLys	499	
QY	337	GTGGGTGGCGCGCTGTAGCTGTGGCTGGCTACAGCTGGGGACGACCTCCTCGAC	396	
DB	500	Val-----CysAspGlyGlnProAspCysLeuAsnGlySerAspGluGluGln	515	
QY	397	TGTCACCCCGCAGTGAAGTTCCCTTGTGGAGGCCCTGGAGCGGATGGAGAAGAGCGC	456	
DB	516	CysGlnGluGlyVal-----ProCysGlyThrPheThrPheGlnCysGlu---AspArg	532	
QY	457	AGTCACCTGAAA-----CGACACACAGAA	480	
DB	533	SerCysValLysLysProAsnProGlnCysAspGlyArgProAspCysArgAspGlySer	552	
QY	481	GACCAGAA-----GACCAAGTAGATCCGGGTCTCATTCATGGGAG	522	
DB	553	AspGluGluHisCysAspCysGlyLeuGlnGlyProSerSerArgIleValGlyGlyAla	572	
QY	523	ATGACACGCGGGGAGACGCCCTCGCAGGTGGTCTGCTGACCTCAAGAAAGAACGCTG	582	
DB	573	ValSerSerGluGlyGluTrpProTrpGln---AlaSerLeuGlnValArgGlyArgHis	591	
QY	583	GCTCGGGGCGAGTCTCATCCACCCCTCTGGGTGTGCACGGGCCGCTCATCGATGGAT	642	
DB	592	IleCysGlyGlyAlaLeuIleAlaAspArgTrpValIleThrAlaIleHisCysPheGln	611	
QY	643	GAGTCCAAGAAGTCTCTGTCAGGCTTGGAGATGATGACCTCGCGCGC---TGGAG---	696	
DB	612	GluAspSerMetAlaSerThrValLeuTrpThrValPheLeuGlyLysValTrpGlnAsn	631	
QY	697	---AAGTGG-----GAGCTGGACCTGGACATCAAGAGGCTTCGTCACCCCAACTAC	747	
		:::		:::
DB	632	SerArgTrpProGlyGluValSerPheLysValSerArgLeuLeuHisProTyrHis	651	
QY	748	AGCAAGAGCACCCAGCAGCAATGACATCGACCTGTGCACCTGGCCCGCCGCCACCTC	807	
		:::		
DB	652	GluGluAspSerHisAspTyrAspValAlaLeuLeuGlnLeuAspHisProValValArg	671	
QY	808	TCGAGACCATAGTGCCTATCTGCTCCCGGACAGCGGCTTCACAGCGCGAGCTCAAT	867	
DB	672	SerAlaAlaValArgProValCysLeuPro-----AlaArgSerHisPhePhe	687	
QY	868	CAGCGCGCGCAGGAGACCTCTGTCAGCGGCTGGGGCTTACCACAGCAGCCGAGAGAGGAG	927	
		:::		
DB	688	GluProGlyLeuHisCysTrpIleThrGlyTrpGly-----AlaLeuArgGluGlyGly	705	
QY	928	GCCAAAGAGAAACCGCACCTCTGCTCCTCACTTCATCAAGATTCGCGTGGTCCGCGACAAT	987	
DB	706	ProIleSerAsn-----AlaLeuGlnLysValAspValGlnLeuIleProGlnAsp	722	
QY	988	GAGTGCAGCGAGGTGATGAGCAACATGTGTGTGAGAACATGTGTGTGTGGCGGATCCTC	1047	
DB	723	LeuCysSerGluAlaTyrArgTyrGlnValThrProArgMetLeuCysAlaGlyTyrArg	742	

1	PRIOR APPLICATION NUMBER: 60/078933
2	PRIOR FILING DATE: 1998-03-20
3	PRIOR APPLICATION NUMBER: 60/078910
4	PRIOR FILING DATE: 1998-03-20
5	PRIOR APPLICATION NUMBER: 60/078939
6	PRIOR FILING DATE: 1998-03-20
7	PRIOR APPLICATION NUMBER: 60/079294
8	PRIOR FILING DATE: 1998-03-25
9	PRIOR APPLICATION NUMBER: 60/079656
10	PRIOR FILING DATE: 1998-03-26
11	PRIOR APPLICATION NUMBER: 60/079664
12	PRIOR FILING DATE: 1998-03-27
13	PRIOR APPLICATION NUMBER: 60/079689
14	PRIOR FILING DATE: 1998-03-27
15	PRIOR APPLICATION NUMBER: 60/079663
16	PRIOR FILING DATE: 1998-03-27
17	PRIOR APPLICATION NUMBER: 60/079728
18	PRIOR FILING DATE: 1998-03-27
19	PRIOR APPLICATION NUMBER: 60/079786
20	PRIOR FILING DATE: 1998-03-27
21	PRIOR APPLICATION NUMBER: 60/079920
22	PRIOR FILING DATE: 1998-03-30
23	PRIOR APPLICATION NUMBER: 60/079923
24	PRIOR FILING DATE: 1998-03-30
25	PRIOR APPLICATION NUMBER: 60/080105
26	PRIOR FILING DATE: 1998-03-31
27	PRIOR APPLICATION NUMBER: 60/080107
28	PRIOR FILING DATE: 1998-03-31
29	PRIOR APPLICATION NUMBER: 60/080165
30	PRIOR FILING DATE: 1998-03-31
31	PRIOR APPLICATION NUMBER: 60/080194
32	PRIOR FILING DATE: 1998-03-31
33	PRIOR APPLICATION NUMBER: 60/080327
34	PRIOR FILING DATE: 1998-04-01
35	PRIOR APPLICATION NUMBER: 60/080328
36	PRIOR FILING DATE: 1998-04-01
37	PRIOR APPLICATION NUMBER: 60/080333
38	PRIOR FILING DATE: 1998-04-01
39	PRIOR APPLICATION NUMBER: 60/080334
40	PRIOR FILING DATE: 1998-04-01
41	PRIOR APPLICATION NUMBER: 60/081070
42	PRIOR FILING DATE: 1998-04-08
43	PRIOR APPLICATION NUMBER: 60/081049
44	PRIOR FILING DATE: 1998-04-08
45	PRIOR APPLICATION NUMBER: 60/081071
46	PRIOR FILING DATE: 1998-04-08
47	PRIOR APPLICATION NUMBER: 60/081195
48	PRIOR FILING DATE: 1998-04-08
49	PRIOR APPLICATION NUMBER: 60/081203
50	PRIOR FILING DATE: 1998-04-09
51	PRIOR APPLICATION NUMBER: 60/081229
52	PRIOR FILING DATE: 1998-04-09
53	PRIOR APPLICATION NUMBER: 60/081955
54	PRIOR FILING DATE: 1998-04-15
55	PRIOR APPLICATION NUMBER: 60/081817
56	PRIOR FILING DATE: 1998-04-15
57	PRIOR APPLICATION NUMBER: 60/081819
58	PRIOR FILING DATE: 1998-04-15
59	PRIOR APPLICATION NUMBER: 60/081952
60	PRIOR FILING DATE: 1998-04-15
61	PRIOR APPLICATION NUMBER: 60/081838
62	PRIOR FILING DATE: 1998-04-15
63	PRIOR APPLICATION NUMBER: 60/082568
64	PRIOR FILING DATE: 1998-04-21
65	PRIOR APPLICATION NUMBER: 60/082569
66	PRIOR FILING DATE: 1998-04-21
67	PRIOR APPLICATION NUMBER: 60/082704
68	PRIOR FILING DATE: 1998-04-22
69	PRIOR APPLICATION NUMBER: 60/082804
70	PRIOR FILING DATE: 1998-04-22
71	PRIOR APPLICATION NUMBER: 60/082700
72	PRIOR FILING DATE: 1998-04-22
73	PRIOR APPLICATION NUMBER: 60/082797

, PRIOR FILING DATE: 1998-04-22  
 , PRIOR APPLICATION NUMBER: 60/082796  
 , PRIOR FILING DATE: 1998-04-23  
 , PRIOR APPLICATION NUMBER: 60/083336  
 , PRIOR FILING DATE: 1998-04-27  
 , PRIOR APPLICATION NUMBER: 60/083322  
 , PRIOR FILING DATE: 1998-04-28  
 , PRIOR APPLICATION NUMBER: 60/083392  
 , PRIOR FILING DATE: 1998-04-29  
 , PRIOR APPLICATION NUMBER: 60/083495  
 , PRIOR FILING DATE: 1998-04-29  
 , PRIOR APPLICATION NUMBER: 60/083496  
 , PRIOR FILING DATE: 1998-04-29  
 , PRIOR APPLICATION NUMBER: 60/083499  
 , PRIOR FILING DATE: 1998-04-29  
 , PRIOR APPLICATION NUMBER: 60/083545  
 , PRIOR FILING DATE: 1998-04-29  
 , PRIOR APPLICATION NUMBER: 60/083554  
 , PRIOR FILING DATE: 1998-04-29  
 , PRIOR APPLICATION NUMBER: 60/083558  
 , PRIOR FILING DATE: 1998-04-29  
 , PRIOR APPLICATION NUMBER: 60/083559  
 , PRIOR FILING DATE: 1998-04-29  
 , PRIOR APPLICATION NUMBER: 60/083500  
 , PRIOR FILING DATE: 1998-04-29  
 , PRIOR APPLICATION NUMBER: 60/083742  
 , PRIOR FILING DATE: 1998-04-30  
 , PRIOR APPLICATION NUMBER: 60/084366  
 , PRIOR FILING DATE: 1998-05-05  
 , PRIOR APPLICATION NUMBER: 60/084414  
 , PRIOR FILING DATE: 1998-05-06  
 , PRIOR APPLICATION NUMBER: 60/084441  
 , PRIOR FILING DATE: 1998-05-06  
 , PRIOR APPLICATION NUMBER: 60/084637  
 , PRIOR FILING DATE: 1998-05-07  
 , PRIOR APPLICATION NUMBER: 60/084639  
 , PRIOR FILING DATE: 1998-05-07  
 , PRIOR APPLICATION NUMBER: 60/084640  
 , PRIOR FILING DATE: 1998-05-07  
 , PRIOR APPLICATION NUMBER: 60/084598  
 , PRIOR FILING DATE: 1998-05-07  
 , PRIOR APPLICATION NUMBER: 60/084600  
 , PRIOR FILING DATE: 1998-05-07  
 , PRIOR APPLICATION NUMBER: 60/084627  
 , PRIOR FILING DATE: 1998-05-07  
 , PRIOR APPLICATION NUMBER: 60/084643  
 , PRIOR FILING DATE: 1998-05-07  
 , PRIOR APPLICATION NUMBER: 60/085339  
 , PRIOR FILING DATE: 1998-05-13  
 , PRIOR APPLICATION NUMBER: 60/085338  
 , PRIOR FILING DATE: 1998-05-13  
 , PRIOR APPLICATION NUMBER: 60/085323  
 , PRIOR FILING DATE: 1998-05-13  
 , PRIOR APPLICATION NUMBER: 60/085582  
 , PRIOR FILING DATE: 1998-05-15  
 , PRIOR APPLICATION NUMBER: 60/085700  
 , PRIOR FILING DATE: 1998-05-15  
 , PRIOR APPLICATION NUMBER: 60/085689  
 , PRIOR FILING DATE: 1998-05-15  
 , PRIOR APPLICATION NUMBER: 60/085579  
 , PRIOR FILING DATE: 1998-05-15  
 , PRIOR APPLICATION NUMBER: 60/085580  
 , PRIOR FILING DATE: 1998-05-15  
 , PRIOR APPLICATION NUMBER: 60/085573  
 , PRIOR FILING DATE: 1998-05-15  
 , PRIOR APPLICATION NUMBER: 60/085704  
 , PRIOR FILING DATE: 1998-05-15  
 , PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:

Pred. No.:

Pred. 1

Score:  
Percent Similarity:

2-070-30

470.00

46.938

11

Length:  
Matghos:

**MacCles:**  
Compassion!

802

124

Best Local Similarity: 33.07% Mismatches: 145  
Query Match: 20.11% Indels: 54  
DB: 9 Gaps: 18

US-09-763-153-1 (1-1245) x US-09-999-832A-169 (1-802)

QY 172 CCGTGGCCAGC-----CTGTGCTGGGGCAGCGCTGCATC-----GACGGC 216  
Db 448 ProCysProGlyGluPheLeuCysSerValAsnGlyLeuCysValProAlaCysAspGly 467  
QY 217 ATCGGCGAGCTCAGCTGCCGCGCAGCGCTGGGAGCGCGCTTCTGCAGCGCGAG 276  
Db 468 Valys-----AspCysProAsnGlyLeuAspGluArgAsnCysValCysArg 483  
QY 277 GTGAGCTTCTCAATTGCTGCGGACACAGCGCGCTCCAGCATCTACTGCTAGAGAG 336  
Db 484 AlaThrPhe---GlnCysLysGluAspSer-----ThrCysIleSerLeuProLys 499  
QY 337 GTGGGCTGGCGGCTGTAGCTGTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 396  
Db 500 Val-----CysAspGlyGlnProAspCysLeuAsnGlySerAspGluGln 515  
QY 397 TGTACCCCGCAGTCAAGTTCCTGTGGGAGCGCTGGAGCGCTGGAGCGGAGGAGAGCGC 456  
Db 516 CysGlnGluGlyVal-----ProCysGlyThrPheThrPheGlnCysGlu---AspArg 532  
QY 457 AGTCACTGAAA-----GACCAAGTAGATCCGGGCTGCTGCTGCTGCTGCTGCTGCT 480  
Db 533 SerCysValLysLysProAsnProGlnCysAspGlyArgProAspCysArgAspGlySer 552  
QY 481 GACCAAGAA-----GACCAAGTAGATCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 522  
Db 553 AspGluGluHisCysAspCysGlyLeuGlnGlyProSerSerArgIleValGlyAla 572  
QY 523 ATGACACGCGCGGAGACAGCCCTGCGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 582  
Db 573 ValSerSerGluGlyGluTrpTrpTrpGln---AlaSerLeuGlnValArgGlyArgHis 591  
QY 583 GCCTGCGGGGAGTGTCTATCCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 642  
Db 592 IleCysGlyGlyAlaLeuIleAlaAspArgTrpValIleThrAlaAlaHisCysPheGln 611  
QY 643 GAGTCCAAAGAGCTCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696  
Db 612 GluAspSerMetAlaSerThrValLeuTrpThrValPheLeuGlyLysValTrpGlnAsn 631  
QY 697 ---AAGTGG-----GAGCTGGACCTGGACATCAAGAGGCTCTGCTGCTGCTGCTGCTGCTGCT 747  
Db 632 SerArgTrpProGlyGluValSerPheLysValSerArgLeuLeuLeuHisProTrpHis 651  
QY 748 AGCAAGAGCACCACGACATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 807  
Db 652 GluGluAspSerHisAspTrpValAlaLeuLeuGlnLeuAspHisProValValArg 671  
QY 808 TCGGAGACCATAGTCCCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 867  
Db 672 SerAlaAlaValArgProValCysLeuPro-----AlaArgSerHisPhePhe 687  
QY 868 CAGCGCGCGCAGAGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 927  
Db 688 GluProGlyLeuHisCysTrpIleThrGlyTrpGly-----AlaLeuArgGluGlyGly 705  
QY 928 GCAAGAGAAACCGCACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 987  
Db 706 ProIleSerAsn-----AlaLeuGlnLysValAspValGlnLeuIleProGlnAsp 722  
QY 988 GAGTGCAGGAGGTATGAGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1047  
Db 723 LeuCysSerGluAlaTrpArgTrpGlnValProArgMetLeuCysAlaGlyTrpArg 742  
QY 1048 GGGGACCGCAGGATGCTGCGAGGCGCAGTGGGGGGCCCATGCTGCTGCTGCTGCTGCTGCTGCT 1104  
Db 743 LysGlyLysLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLysAlaLeu 762

QY 1105 CACGCGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1164  
Db 763 SerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArgProAsn 782  
QY 1165 AACTACGGGCTTTACACCAAGTACGACCGCTACCTCGACTGGATC 1209  
Db 783 TyrPheGlyValTyThrArgIleThrGlyValIleSerTrpIle 797

## RESULT 17

US-09-978-189-169  
; Sequence 169, Application US/09978189  
; Publication No. US20030004102A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kijavlin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2630PIC7  
; CURRENT APPLICATION NUMBER: US/09/978,189  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 09/918585  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/064249  
; PRIOR FILING DATE: 1997-11-03  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066364  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/077450  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: 60/077632  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077641  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077649  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077791  
; PRIOR FILING DATE: 1998-03-12  
; PRIOR APPLICATION NUMBER: 60/078004  
; PRIOR FILING DATE: 1998-03-13  
; PRIOR APPLICATION NUMBER: 60/078886  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078936  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078910

us-09-763-153-1.n2p.rapb

Tue Mar 18 16:19:31 2003

```

; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

```

Alignment Scores:  
Pred. No.: 2.07e-20  
Score: 470.00  
Percent Similarity: 46.93%  
Best Local Similarity: 33.07%  
Query Match: 20.11%  
DB: 9

Length:  
Matches: 802  
Conservative: 124  
Mismatch: 52  
Indels: 145  
Gaps: 54  
18

US-09-763-153-1 (1-1245) x US-09-978-189-169 (1-802)

QY 172 CCGTGGCCAGC-----CTGTGCTGGGCGACGGCGGCTGTC-----GACGGC 216  
 Db 448 ProCysProGlyGluPheLeuLeuValSerValAsnGlyLeuCysValProAlaCysAspGly 467

QY 217 ATCGCGAGCTTACGTCGACTCCGCGAGCGGCTGGGCGGCTCTGCGCAGCGGAG 276  
 Db 468 ValLys-----AspCysProAsnGlyLeuAspGluArgAsnCysValCysArg 483

QY 277 GTGAGCTTCTCAATTTGCTGCGTGGCAACGCGGCTGCACGCATTACTGCGCAGCGGAG 336  
 Db 484 AlaThrPhe---GlnCysLysGluAspSer-----ThrCysIleSerLeuProLys 499

QY 337 GTGGGCTGGCGGCGCTGTAGCTGCGCGCTGCTACAACTGGGCGGAGAGCCTCTCTGAG 396  
 Db 500 Val-----CysAspGlyGlnProAspCysLeuAsnGlySerAspGluGln 515

QY 397 TGTACACCGCAGTGAAGTTCCCTGTGGAGCGCCTGGAAGCGGATGGAGAAGAGCGC 456  
 Db 516 CysGlnGluGlyVal-----ProCysGlyThrPheThrPheGlnCysGlu---AspArg 532

QY 457 AGTCACCTGAAA-----GACCAAGTAGATCCGCGCTCATGTGGGAAG 522  
 Db 533 SerCysValLysLysProAsnProGlnCysAspGlyArgProAspCysArgAspGlySer 552

QY 481 GACCAAGAA-----GACCAAGTAGATCCGCGCTCATGTGGGAAG 522  
 Db 553 AspGluGlnHisCysAspCysGlyLeuGlnGlyProSerSerArgIleValGlyGlyAla 572

QY 523 ATGACGAGCGGGGAGACAGCCCTGCGAGCTGTCTGCTGGACTCAAGAAGAAGCTG 582  
 Db 573 ValSerSerGluGlyGluTrpTrpGln---AlaSerLeuGlnValArgGlyArgHis 591

QY 583 GCGTCGGGGGAGTCTCATCCACCCCTCTGCTGGTGTGACAGCGGCCCTCATGGAT 642  
 Db 592 IleCysGlyGlyAlaLeuIleAlaAspArgTrpValIleThrAlaAlaHisCysPheGln 611

QY 643 GAGTCCAAAGAGCTCTTGTGAGCTGGAGATGACCTGGCGGC---TGGAG--- 696  
 Db 612 GluAspSerMetAlaSerThrValLeuTrpThrValPheLeuGlyLysValTrpGlnAsn 631

QY 697 ---AAGTGG-----GAGCTGGAGCTGGACATCAAGGAGCTCTGCTCCACCCCACTAC 747  
 Db 632 SerArgTrpProGlyGluValSerPheLysValSerArgLeuLeuLeuHisProTyrHis 651

QY 748 AGCAAGACACCCAGACATGACATCGACTGCTGACCTGCCCGCCAGCCGCCCTC 807  
 Db 652 GluGluAspSerHisAspTyrAspValAlaLeuLeuGlnLeuAspHisProValValArg 671

QY 808 TCGCAGACCATGATGCGCATCTGCTGCTCCCGCAGCAGCGGCTGCGAGCGCGCTCAAT 867  
 Db 672 SerAlaAlaValArgProValCysLeuPro-----AlaArgSerHisPhePhe 687

QY 868 CAGCGCGCGCAGGAGCCCTGTCGAGCGGCTGGGGTACACAGCAGCGCGAGAGAAGAG 927  
 Db 688 GluProGlyLeuHisCysTrpIleThrGlyTrpGly-----AlaLeuArgGluGlyGly 705

QY 928 GCCAAGAGAAACCGCACCTCTGCTGCTCAACTCATCAAGATTCCCGTGGTCCGCACAA 987  
 Db 706 ProfileSerAsn-----AlaLeuGlnLysValAspValGlnLeuIleProGlnAsp 722

QY 988 GAGTCGACGAGGTATGAGCAACATGTTCTGAGAATCTGCTGTCGGGGATCCTC 1047  
 Db 723 LeuCysSerGluAlaTyrArgTyrGlnValThrProArgMetLeuCysAlaGlyTyrArg 742

QY 1048 GGGAGCCGCGAGGATCCCTGCGAGCGCAGACTGGGGGGCCCATGTGCGCC---TCCTTC 1104  
 Db 743 LysGlyLysLysAspLysGlnGlyAspSerGlyGlyProLeuValCysLysAlaLeu 762

QY 1105 CAGCGCACCTGTTCTTGTGGGCGCTGTGAGCTGGGCTGAGGCTGTGGCTCTTCAC 1164  
 Db 762 ThrCysIleSerLeuProLys 499

Db 763 SerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArgProAsn 789

QY 1165 AACTACCGCGCTTTACACCAAGTACGCGCTACTCTGCTGGATC 1209  
 Db 783 TyrPheGlyValTyrThrArgIleThrGlyValIleSerTrpIle 797

RESULT 18  
 US-09-978-608A-169  
 ; Sequence 169, Application US/09978608A  
 ; Publication No. US20030045462A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Baker Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Ferrara, Napoleon  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Hillan, Kenneth J.  
 ; APPLICANT: Kljavin, Ivar J.  
 ; APPLICANT: Kuo, Sophia S.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James;  
 ; APPLICANT: Paoni, Nicholas F.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Shelton, David L.  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; TITLE OF INVENTION: Acids Encoding the Same  
 ; FILE REFERENCE: P2630P1C22  
 ; CURRENT APPLICATION NUMBER: US/09/978,608A  
 ; CURRENT FILING DATE: 2001-10-16  
 ; NUMBER OF SEQ ID NOS: 624  
 ; Prior Application removed - See File Wrapper or Palm  
 ; SEQ ID NO 169  
 ; LENGTH: 802  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-978-608A-169

Alignment Scores:  
 Pred. No.: 2,07e-20 Length: 802  
 Score: 470.00 Matches: 124  
 Percent Similarity: 46.93% Conservative: 52  
 Best Local Similarity: 33.07% Mismatches: 145  
 Query Match: 20.11% Indels: 54  
 Db: 9 Gaps: 18

US-09-763-153-1 (1-1245) x US-09-978-608A-169 (1-802)

QY 172 CCGTGGCCAGC-----CTGTGCTGGGCGACGGCGGCTGTC-----GACGGC 216  
 Db 448 ProCysProGlyGluPheLeuLeuValSerValAsnGlyLeuCysValProAlaCysAspGly 467

QY 217 ATCGCGAGCTTACGTCGACTCCGCGAGCGGCTGGGCGGCTCTGCGCAGCGGAG 276  
 Db 468 ValLys-----AspCysProAsnGlyLeuAspGluArgAsnCysValCysArg 483

QY 277 GTGAGCTTCTCAATTTGCTGCGTGGCAACGCGGCTGCACGCATTACTGCGCAGCGGAG 336  
 Db 484 AlaThrPhe---GlnCysLysGluAspSer-----ThrCysIleSerLeuProLys 499

QY 337 GTGGGCTGGCGGCGCTGTAGCTGCGCGCTGCTACAACTGGGCGGAGAGCCTCTCTGAG 396  
 Db 500 Val-----CysAspGlyGlnProAspCysLeuAsnGlySerAspGluGln 515

QY 397 TGTACACCGCAGTGAAGTTCCCTGTGGAGCGCCTGGAAGCGGATGGAGAAGAGCGC 456  
 Db 516 CysGlnGluGlyVal-----ProCysGlyThrPheThrPheGlnCysGlu---AspArg 532

QY 457 AGTCACCTGAAA-----GACCAAGTAGATCCGCGCTCATGTGGGAAG 522  
 Db 533 SerCysValLysLysProAsnProGlnCysAspGlyArgProAspCysArgAspGlySer 552

QY 481 GACCAAGAA-----GACCAAGTAGATCCGCGCTCATGTGGGAAG 522  
 Db 553 AspGluGlnHisCysAspCysGlyLeuGlnGlyProSerSerArgIleValGlyGlyAla 572

QY 523 ATGACGAGCGGGGAGACAGCCCTGCGAGCTGTCTGCTGGACTCAAGAAGAAGCTG 582  
 Db 573 ValSerSerGluGlyGluTrpTrpGln---AlaSerLeuGlnValArgGlyArgHis 591

QY 583 GCGTCGGGGGAGTCTCATCCACCCCTCTGCTGGTGTGACAGCGGCCCTCATGGAT 642  
 Db 592 IleCysGlyGlyAlaLeuIleAlaAspArgTrpValIleThrAlaAlaHisCysPheGln 611

QY 643 GAGTCCAAAGAGCTCTTGTGAGCTGGAGATGACCTGGCGGC---TGGAG--- 696  
 Db 612 GluAspSerMetAlaSerThrValLeuTrpThrValPheLeuGlyLysValTrpGlnAsn 631

QY 697 ---AAGTGG-----GAGCTGGAGCTGGACATCAAGGAGCTCTGCTCCACCCCACTAC 747  
 Db 632 SerArgTrpProGlyGluValSerPheLysValSerArgLeuLeuLeuHisProTyrHis 651

QY 748 AGCAAGACACCCAGACATGACATCGACTGCTGACCTGCCCGCCAGCCGCCCTC 807  
 Db 652 GluGluAspSerHisAspTyrAspValAlaLeuLeuGlnLeuAspHisProValValArg 671

QY 808 TCGCAGACCATGATGCGCATCTGCTGCTCCCGCAGCAGCGGCTGCGAGCGCGCTCAAT 867  
 Db 672 SerAlaAlaValArgProValCysLeuPro-----AlaArgSerHisPhePhe 687

QY 868 CAGCGCGCGCAGGAGCCCTGTCGAGCGGCTGGGGTACACAGCAGCGCGAGAGAAGAG 927  
 Db 688 GluProGlyLeuHisCysTrpIleThrGlyTrpGly-----AlaLeuArgGluGlyGly 705

QY 928 GCCAAGAGAAACCGCACCTCTGCTGCTCAACTCATCAAGATTCCCGTGGTCCGCACAA 987  
 Db 706 ProfileSerAsn-----AlaLeuGlnLysValAspValGlnLeuIleProGlnAsp 722

QY 988 GAGTCGACGAGGTATGAGCAACATGTTCTGAGAATCTGCTGTCGGGGATCCTC 1047  
 Db 723 LeuCysSerGluAlaTyrArgTyrGlnValThrProArgMetLeuCysAlaGlyTyrArg 742

QY 1048 GGGAGCCGCGAGGATCCCTGCGAGCGCAGACTGGGGGGCCCATGTGCGCC---TCCTTC 1

```

337 GTGGGCTGGCGGCTGTAGCTGTGGCTGTACAGCTGGGAGCAGCCTCTCGAG 396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
500 Val-----CysAspGlyGlnProAspCysLeuAsnGlySerAspGluGln 515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
397 TGTACCCCGCAGTGAAGTCCCTGTGGAGGCCCTGGAAGCGATGGAGAGAGCCG 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
516 CysGlnGluGlyVal-----ProCysGlyThrPheThrPheGlnCysGlu---AspArg 532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
457 AGTCACCTGAAA-----CGAGACACAGAA 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
533 SerCysValLysProAsnProGlnCysAspGlyArgProAspCysArgAspGlySer 552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
481 GACCAAGAA-----GACCAAGTAGATCCCGGCTCATTTGATGGGAG 522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
553 AspGluHisCysAspCysGlyLeuGlnGlyProSerArgIleValGlyGlyAla 572
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
523 ATGACACGGGGGAGACACCCCTGGCAGGTGGTCTCTGCTGACTCAAGAAGAGCTG 582
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
573 ValSerSerGluGlyGluTrpProTrpGln---AlaSerLeuGlnValArgGlyHis 591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
583 GCCTGGCGGCGAGTGTCTATCCACCCCTCTGGTGTGTGACAGCGGCCCACTGCTGAT 642
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
592 IleCysGlyGlyAlaLeuIleAlaAspArgTrpValIleThrAlaAlaHisCysPheGln 611
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
643 GAGTCCAAGAGCTCTTGTGAGCTGGAGAGTATGACCTGGCGGC---TGGAG--- 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
612 GluAspSerMetAlaSerThrValLeuTrpThrValPheLeuGlyLysValTrpGlnAsn 631
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
697 ---AAGTGG-----GAGCTGGACCTGGACATCAAGAGGCTCTTCTGCCACCCCACTAC 747
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
632 SerArgTrpProGlyGluValSerPheLysValSerArgLeuLeuHisProTrpHis 651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
748 AGCAAGAGCACCACCGACATGACATCGCATGCTGACCTGCCCGCCCGCCCACTAC 807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
652 GluGluAspSerHisAspTyAspValAlaLeuLeuGlnLeuAspHisProValValArg 671
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
808 TCGAGACCATAGTCCCATCTGCTCCGACAGCGGCTGTGACAGCGCGAGCTCAAT 867
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
672 SerAlaAlaValArgProValCysLeuPro-----AlaArgSerHisPhePhe 687
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
868 CAGCGCGCGGAGAGACCTCGTACGCGGCTGGGCTACACAGCAGCGAGAGAGGAG 927
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
688 GluProGlyLeuHisCysTrpIleThrGlyTrpGly-----AlaLeuArgGluGly 705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
928 GCCAAGAGAAACCGACCTCTGCTCACTTCAAGATTCCCGTGGTCCCGCACAAT 987
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
706 ProIleSerAsn-----AlaLeuGlnLysValAspValGlnLeuIleProGlnAsp 722
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
988 GAGTGCAGCGAGTCAATGACCAACATGCTGTGTGAGAACATGCTGTGGCGGCATCTC 1047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
723 LeuCysSerGluAlaTrpArgTrpGlnValThrProArgMetLeuCysAlaGlyTrpArg 742
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1048 GGGGACCGGAGGATCGCTGGAGGCGACAGTGGGGGCCCGCTGGTGGCC---TCCTTC 1104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
743 LysGlyLysLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLysAlaLeu 762
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1105 CACGACACCTGTCTGCTGGTGGCTGTGAGCTGGGCTGGAGGCTGTGGGCTCTCTCAC 1164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
763 SerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArgProAsn 782
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1165 AACTACGGCTTTACACCAAGTACCGCCCTACCTCGACTGGATC 1209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
783 TyrPheGlyValTrpThrArgIleThrGlyValIleSerTrpIle 797
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 19

```

US-09-888-615-113
; Sequence 113, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARVDCZAK, GLEN

```

```

; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-888-615-113

```

```

Alignment Scores:
pred. No.: 2,07e-20 Length: 802
Score: 470.00 Matches: 124
Percent Similarity: 46.93% Conservative: 52
Best Local Similarity: 33.07% Mismatches: 145
Query Match: 20.11% Indels: 54
DB: 10 Gaps: 18

```

US-09-763-153-1 (1-1245) x US-09-888-615-113 (1-802)

```

QY 172 CCGTGGCCGACG-----CTGTGCTGGGGCAGCGGACGCTGCATC-----GACGGC 216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 448 ProCysProGlyGluPheLeuCysSerValAsnGlyLeuCysValProAlaCysAspGly 467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 217 ATCGGCGAGTTCAGCTCGACTCCGCGAGCGGCTGGAGGCGGCTGTGCGCAGCGCGAG 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 468 ValLys-----AspCysProAsnGlyLeuAspGluArgAsnGlyValCysArg 483
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 277 GTGAGCTTCCTCAATTGCTGTGGACACAGCGGCTGACCGATTACTGCTAGAGAG 336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 484 AlaThrPhe-----GlnCysLysGluAspSer-----ThrCysIleSerLeuProLys 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 337 GTGGCTGGCGGCGCTAGTGTGCGCTGGCTGCTACAGCTGGGGAGCAGCTCTCTGCGAG 396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 500 Val-----CysAspGlyGlnProAspCysLeuAsnGlySerAspGluGln 515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 397 TGTACCCCGCAGTGAAGTCCCTGTGGAGGCCCTGGAAGCGGATGGAGAGAGAGCCG 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 516 CysGlnGluGlyVal-----ProCysGlyThrPheThrPheGlnCysGlu---AspArg 532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 457 AGTCACCTGAAA-----CGAGACACAGAA 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 533 SerCysValLysProAsnProGlnCysAspGlyArgProAspCysArgAspGlySer 552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 481 GACCAAGAA-----GACCAAGTAGATCCCGGCTCATTTGATGGGAG 522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 553 AspGluHisCysAspCysGlyLeuGlnGlyProSerArgIleValGlyGlyAla 572
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 523 ATGACACAGCGGGGAGACACCCCTGGCAGGTGGTCTCTGCTGACTCAAGAAGAGAGCTG 582
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 573 ValSerSerGluGlyGluTrpProTrpGln---AlaSerLeuGlnValArgGlyHis 591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 583 GCCTGGCGGCGAGTGTCTATCCACCCCTCTGGTGTGTGACAGCGGCCCACTGCTGATGAT 642
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 592 IleCysGlyGlyAlaLeuIleAlaAspArgTrpValIleThrAlaAlaHisCysPheGln 611
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 643 GAGTCCAAGAGCTCTTGTGAGCTGGAGAGTATGACCTGGCGGC---TGGAG--- 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 612 GluAspSerMetAlaSerThrValLeuTrpThrValPheLeuGlyLysValTrpGlnAsn 631
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 697 ---AAGTGG-----GAGCTGGACCTGGACATCAAGAGGCTCTTCTGCCACCCCACTAC 747
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 632 SerArgTrpProGlyGluValSerPheLysValSerArgLeuLeuHisProTrpHis 651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 748 AGCAAGAGCACCACCGACATGACATCGCATGCTGACCTGCCCGCCCGCCCACTAC 807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 652 GluGluAspSerHisAspTyAspValAlaLeuLeuGlnLeuAspHisProValValArg 671
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 808 TCGAGACCATAGTCCCATCTGCTCCGACAGCGGCTGTGACAGCGCGAGCTCAAT 867
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 672 SerAlaAlaValArgProValCysLeuPro-----AlaArgSerHisPhePhe 687
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 868 CAGCGCGCGGAGAGACCTCGTACGCGGCTGGGCTACACAGCAGCGAGAGAGGAG 927
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 688 GluProGlyLeuHisCysTrpIleThrGlyTrpGly-----AlaLeuArgGluGly 705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 928 GCCAAGAGAAACCGACCTCTGCTCACTTCAAGATTCCCGTGGTCCCGCACAAT 987
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 706 ProIleSerAsn-----AlaLeuGlnLysValAspValGlnLeuIleProGlnAsp 722
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 988 GAGTGCAGCGAGTCAATGACCAACATGCTGTGTGAGAACATGCTGTGGCGGCATCTC 1047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 723 LeuCysSerGluAlaTrpArgTrpGlnValThrProArgMetLeuCysAlaGlyTrpArg 742
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1048 GGGGACCGGAGGATCGCTGGAGGCGACAGTGGGGGCCCGCTGGTGGCC---TCCTTC 1104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 743 LysGlyLysLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLysAlaLeu 762
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1105 CACGACACCTGTCTGCTGGTGGCTGTGAGCTGGGCTGGAGGCTGTGGGCTCTCTCAC 1164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 763 SerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArgProAsn 782
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1165 AACTACGGCTTTACACCAAGTACCGCCCTACCTCGACTGGATC 1209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 783 TyrPheGlyValTrpThrArgIleThrGlyValIleSerTrpIle 797
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



QY	808	TCGCAGACCAATAGTGGCCACTCTGCCCTCCCGGACAGCGCGCTTGCCAGACGGGAGCTCAAT	867
Db	672	SerAlaAlaValArgProValCysLeuPro-----AlaArgSerHisPhePhe	687
QY	868	CAGSCCGGCCAGGAGACCCCTCGTGTAGCGGCTGGGCTACCCAGACAGCCGAGAGAAGAG	927
Db	688	GluProGlyLeuHisCysTrpIleThrGlyTrpGly-----AlaLeuArgGluGlyGly	705
QY	928	GCCAAGAAACACCGCCACCTTCGTCTCACTTCACTCAAGATTCCCGTGGTCCCGCACAAAT	987
Db	706	ProfileSerAsn-----AlaLeuGlnLysValAspValGlnLeuIleProGlnAsp	722
QY	988	GAGTGCAGCAGGCTCATGAGCAACATGGTGTCTGAGACATCTGTGTGGGGGCATCCTC	1047
Db	723	LeuCysSerGluAlaTyrArgTyrGlnValTrpProArgMetLeuCysAlaGlyTyrArg	742
QY	1048	GGGACCGCAGGATCCCTCGCAGGCGACAGTGGGGGCCCATCGTGGCC-----TCCTTC	1104
Db	743	LysGlyLysLysAspLacysGlnGlyaspSerGlyGlyProLeuValCysLysAlaLeu	762
QY	1105	CACGGCACCTGGTCTCTGGTGGCCCTGTGAGCTGGGTGAGGGCTGTGGGCTCCTTCAC	1164
Db	763	SerGlyArgtrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArgProAsn	782
QY	1165	AACCTACGGCTTACACCAAGATCAGCCGCTACCTCGACTGGATC	1209
Db	783	TyrPheGlyValTyrThrArgIleThrGlyValIleSerTrpIle	797
RESULT 20			
US-09-888-615-108			
; Sequence 108, Application US/09888615			
; Patent No. US20020064856A1			
; GENERAL INFORMATION:			
; APPLICANT: PLOWMAN, GREGORY			
; APPLICANT: WHYTE, DAVID			
; APPLICANT: CAENEPEEL, SEAN			
; APPLICANT: CHARYDCZAK, GLEN			
; APPLICANT: MANNING, GERARD			
; APPLICANT: SUDARSANAM, SUCHA			
; TITLE OF INVENTION: NOVEL PROTEASES			
; FILE REFERENCE: 038602/1214			
; CURRENT APPLICATION NUMBER: US/09/888,615			
; CURRENT FILING DATE: 2001-06-26			
; PRIOR APPLICATION NUMBER: 60/214,047			
; PRIOR FILING DATE: 2000-06-26			
; NUMBER OF SEQ ID NOS: 150			
SOFTWARE: Patent In Ver. 2.1			
SEQ ID NO 108			
LENGTH: 850			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-888-615-108			
Alignment Scores:			
Pred. No.:			
Score: 5.62e-18 Length: 850			
Best Local Similarity: 43.99% Matches: 120			
Query Match: 30.69% Mismatches: 52			
DB: 18.34% Indels: 69			
Gaps: 16			
US-09-763-1 (1-1245) x US-09-888-615-108 (1-850)			
QY	163	TTGGAGCACCCGTGC-----GCCAGCTGTGTCTGGGGCAGCGCAGCTGCATC	210
Db	487	IleSerGlnProCysProValGlySerPheArgCysSerSerGlyLeuCysValProGln	506
QY	211	-----GACCGCATCGGACGCTTCAGCTGCACATCCCGCAGCGCTGGGAGGC	258
Db	507	AlaGlnArgCysAspGlyValasn-----AspCysPheAspGluSerAspGlu	522
QY	259	CGCTTCTGCCAGCGGAGGTGAGCTCTCCTCAATTGCTCG-----	297

Tue Mar 18 16:19:31 2003

us-09-763-153-1.n2p.rapb

Page 24

Job time : 66 secs